An Outline of Algebraic Set Theory with a View Towards Cohen's Model Falsifying the Continuum Hypothesis

MSc Thesis *(Afstudeerscriptie)*

written by

Julia Ilin
(born November 8th, 1986 in Bremen, Germany)

under the supervision of **Dr Benno van den Berg** and **Prof Dr Benedikt Löwe**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
January 28, 2014

Members of the Thesis Committee:
Dr Benno van den Berg
Dr Nick Bezhanishvili
Dr Yurii Khomskii
Dr Kohei Kishida
Prof Dr Benedikt Löwe
Dr Jakub Szymanik (Chair)

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis is concerned with the area of algebraic set theory. Algebraic set theory was invented by Joyal and Moerdijk [18] with the aim to study set theory from the perspective of category theory. The central notion is a category of classes, given by a triple $(\mathcal{E}, \mathcal{S}, \mathcal{P}_\mathrm{s})$, consisting of a Heyting pretopos $\mathcal{E}$, a particular class $\mathcal{S}$ of arrows of $\mathcal{E}$ that are called small maps and an endofunctor $\mathcal{P}_\mathrm{s} : \mathcal{E} \to \mathcal{E}$. The small maps provide an abstract notion of smallness on $\mathcal{E}$, whereas the endofunctor $\mathcal{P}_\mathrm{s}$ should be thought of as generalized powerclass functor. Universes of set theory arise as initial algebras for this functor.

The main goal of this thesis is prove that Cohen's model negating the continuum hypothesis can be recovered in the algebraic set theory framework. Cohen's model has already been examined in the filed of topos theory by Tierney [29]. It will be shown that Tierney's proof translates to the algebraic set theory setting.

Contents

Chapter 1

Introduction

Algebraic set theory aims to study (intuitionistic) Zermelo Fraenkel set theory from the perspective of category theory. The framework was invented by André Joyal and Ieke Moerdijk and first presented in their book [18]. The main idea of the setting can be summarized as follows. Given a Heyting pretopos $\mathcal{E}$ together with a particular collection of arrows of $\mathcal{E}$ that we call small maps, one can define a generalized powerclass functor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$ on $\mathcal{E}$. Joyal and Moerdijk show in [18] that initial algebras for this functor satisfy the axioms of Zermelo Fraenkel set theory. In this light set theory is formalized by operations on sets rather than by properties of the membership relation and therefore obtains an algebraic flavor.

Before we examine the abstract categorical setting more closely, let us first see how usual set theory fits into this framework. Assume that V is a universe of set theory. Consider the category $\mathcal{E}$ that has the subclasses of V as objects and (possibly class-sized) functions between these classes as arrows. Define a powerclass functor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$ on $\mathcal{E}$ by mapping a class X to the class of all its subsets. An *algebra for the functor* $\mathcal{P}_{\mathrm{s}}$ is a class A, together with a map $\mathcal{P}_{\mathrm{s}}(A) \to A$. Such an algebra $\eta : \mathcal{P}_{\mathrm{s}}(V) \to V$ is called *initial*, if for any algebra $\alpha : \mathcal{P}_{\mathrm{s}}(A) \to A$, there is a unique map $f : V \to A$ such that the diagram

$$\begin{array}{ccc} \mathcal{P}_{\mathrm{s}}(V) & \xrightarrow{\mathcal{P}_{\mathrm{s}}(f)} & \mathcal{P}_{\mathrm{s}}(A) \\ {\scriptstyle\eta}\downarrow & & \downarrow{\scriptstyle\alpha} \\ V & \xrightarrow{\quad f \quad} & A \end{array}$$

commutes. In our setting, the set theoretic universe V that we started with becomes an algebra for the $\mathcal{P}_{\mathrm{s}}$-functor by sending a subset a of V to a itself. The recursion theorem familiar from set theory tells us that our universe V is in fact the initial algebra for the powerclass functor $\mathcal{P}_{\mathrm{s}}$. A central observation made in algebraic set theory is that the abstract property of being an initial algebra for such a powerclass functor is in some sense sufficient to be a universe of set theory. In the framework of algebraic set theory the ingredients of our example are replaced with notions from category theory. Roughly, this can be illustrated as in the following table:

Set theoretic notion	Categorical notion
The true category of classes	Heyting pretopos
Sets	Small objects determined by small maps
Powerclass functor	$\mathcal{P}_{\mathrm{s}}$ functor
Universe of set theory	Initial algebra for the $\mathcal{P}_{\mathrm{s}}$ functor

In the general setting, the true category classes is replaced by the notion of a Heyting pretopos. A Heyting pretopos is a category having similar closure properties as the true category of classes. In particular, it admits elementary set theoretic operations such as products, unions and pairs. Following our example again, we think of the objects of a Heyting pretopos as classes and of its arrows as functions between classes. As in our example we would like to recognize those elements of the Heyting pretops that are small, i.e. that represent sets in some abstract sense. To this end the Heyting pretopos $\mathcal{E}$ is equipped with a class $\mathcal{S}$, a collection of arrows of $\mathcal{E}$ satisfying certain axioms that we call *small maps*. The axioms for the small maps intend to determine those maps as small that have set-sized fibres. In this way the small maps also provide an abstract notion of smallness for objects. The small objects of $\mathcal{E}$ are the ones whose unique map to the terminal object is small.

The axioms for the small maps imply that there is an endofunctor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$ that we call the generalized powerclass functor. Just as the powerclass functor in our example, the functor $\mathcal{P}_{\mathrm{s}}$ intuitively maps an object of $\mathcal{E}$ to the collection of all its small subobjects. Now universes of set theory in the abstract setting arise as initial algebras for the $\mathcal{P}_{\mathrm{s}}$ functor. We call them *algebraic universes*.

The universes occuring in algebraic set theory satisfy in general some intuitionistic form of set theory. In fact, it is one great feature of the algebraic set theory framework that it allows one to obtain models for various set theories in a uniform way. Roughly, this is due to two (interrelated) parameters. On the one hand, the internal logic of the base category $\mathcal{E}$ determines whether the resulting model of set theory satisfies intuitionistic or classical rules. On the other hand, changing the abstract notion of smallness by modifying the axioms for the class of small maps allows one to obtain models of several intuitionistic and constructive set theories. One line of research focuses on isolating axioms for the class $\mathcal{S}$ of small maps to obtain models for a particular set theory. In the original approach [18] soundness with respect to Friedman's **IZF** was proven. Simpson [28] considered a much reduced axiomatization compared to the original framework and proved soundness for the theory **IZF**, but with replacement instead of the collection axiom. He also proves a completeness result, an improvement compared the original approach. Van den Berg and Moerdijk [5] examine Aczel's theory **CZF** in the algebraic set theory setting, also obtaining a completeness result. Recovering models of a particular set theory in this way can be used to lift results from topos theory to the set theory in question. Examples of this kind can be found in [8] for the set theory **CZF**.

Another line of research in algebraic set theory asks whether familiar results from topos theory can be translated into categories of classes. In a series of papers [6, 7], Van den Berg and Moerdijk show that categories of classes are closed under the formation of sheaves and realizability models. These observations allow one in particular to develop sheaf models in the algebraic set theory setting without reference to some preexisting universe of sets.

A different approach is taken in [26, 2]. Instead of examining an established constructive set theory, the authors develop the set theory **BIST**. The theory **BIST** intends to capture the set theory internal to a topos. In particular, the authors show that in fact every topos occurs as the category of sets in a particular category of classes by building the category of ideals of a topos. Also predicative versions of the set theory **BIST** were studied [3].

1.1 Historical Development

We would like to summarize some of the major steps that led to the development of algebraic set theory since the invention of category theory by Eilenberg and Mac Lane in 1945.

In the 1960s, Lawvere and Tierney developed the notion of an elementary topos. Even though the axiomatization of a topos is very simple, it allows almost all mathematical constructions to be carried out internally. In fact, Lawvere's vision was that a topos can be seen as a universe of sets and in this way serves as a foundation for mathematics alternative to set theory. This idea was also strengthened as in the 1970s Mitchell and Bénabou defined a formal language internal to a topos. The internal language allows one to reason about a topos as if it would consist of sets and functions between these. One fascinating aspect of the internal logic of a topos is that it follows intuitionistic rather than classical rules. This brought back the ideas of Brouwer and Heyting developed in the beginning of the 20th century.

Around the same time, a major development in set theory was Cohen's invention of the method of forcing [11], establishing the independence of the continuum hypothesis and the independence of the axiom of choice from the other Zermelo Fraenkel axioms of set theory. Forcing soon became an indispensable proof technique, yielding answers to questions that have long been open by proving them to be independent of the set theoretic axioms. In [29], Tierney presented a categorical version of Cohen's forcing model falsifying the continuum hypothesis. His work was extended by Bunge [10], who defined a topos falsifying Souslin's hypothesis. A bit later, Freyd [14] found a surprisingly simple construction of a topos falsifying the axiom of choice.

These results aroused interest to determine the precise relationship between axiomatic set theory and topos theory. Accordingly, there have been attempts to find models for (intuitionistic) axiomatic set theories inside a topos. In the 1980s Fourman [12] and independently Hayashi [16] suggested methods to interpret set theory in Grothendiek toposes. Also, Scott [27] worked on interpretations of set theory in presheaf categories, presented his work in talks but never published it. The common idea of these models was to mimic the construction of a cumulative hierarchy in a topos. These constructions, however, had two disadvantages. A first minor disadvantage is that the construction proceeds over an external induction over the ordinals. Therefore, the construction is somewhat non-elementary as it needs a reference to set theory and cannot be carried out in the topos itself. Another more serious problem is that quantifiers in the internal logic of a topos are necessarily bounded by objects of the topos. Therefore, to formulate unbounded axioms such as replacement or unrestricted separation, one has to extend the language of set theory with predicates for each object occurring in the cumulative hierarchy.

In the framework of algebraic set theory these problems are overcome. In a Heyting pretopos, classes become the first-class citizens as opposed to sets in topos theory. In this way, the universe of set theory now becomes a particular object of the Heyting pretopos. This object is used to interpret the free variables of the language of set theory. As the bound on variables is now given by the object that is the universe itself, the bound does not cause a restriction anymore. In this light, algebraic set theory can be seen as a complete answer to the correspondence between axiomatic (intuitionistic) set theory and topos theory.

1.2 Outline of the Thesis

The main goal of the present thesis is to show how Cohen's model falsifying **CH** fits into the framework of algebraic set theory. In [29], Tierney has shown that Cohen's model can be recovered in the categorical context by building sheaves over the Cohen poset, a major achievement in topos theory. In the thesis, we will see that Tierney's result fits in the framework of algebraic set theory. To this end we will consider the category of an enlarged version of sheaves over the Cohen poset that form a category of classes. As in [29], one can show that the continuum hypothesis is falsified in the internal logic of this category of classes. Our contribution is to show that the statement in the internal logic of the outer category of classes translates to a statement about the algebraic universe that the category of classes contains. Accordingly, we will obtain an algebraic model of Zermelo Fraenkel set theory that falsifies the continuum hypothesis.

As we explained, earlier topos theory failed to provide a model satisfying all the axioms of Zermelo Fraenkel set theory whereas algebraic set theory succeeds to provide such universes. In this light, our result can be seen as completing the task of recovering Cohen's result in a categorical context. It should be noted, however, that the idea of building a cumulative hierarchy inside a topos, as suggested in [12] or [16], leads to a very similar result.

The major part of the thesis is intended to introduce the reader to the field of algebraic set theory. We will develop the necessary background from category theory that is needed to understand basic concepts and theorems of this field. In particular, we will discuss how to interpret logic inside a category of classes and prove that for our axiomatization the algebraic universes satisfy the set theory **IZF**. Often we will focus on providing detailed proofs of statements that are omitted in the literature.

Chapter 2 In this chapter we will introduce the reader to the notion of a category of classes. We will define the notion of a Heyting pretopos and summarize some of its elementary properties. In particular, we will examine how Heyting pretoposes provide sound semantics with respect to first order intutionistic logic. Moreover, we will provide an axiomatization of the class of small maps leading to the notion of a category of classes. In the final part we will discuss more properties of categories of classes that are used later in the thesis.

Chapter 3 In the third chapter, we will examine examples of categories of classes. First, we will discuss the example of the true category of classes again. Then we will investigate categories of large presheaves and large sheaves. These categories can be seen as a class-sized version of presheaves and sheaves familiar from topos theory.

Chapter 4 In the fourth chapter we will investigate algebraic universes of set theory, i.e. initial algebras for the $\mathcal{P}_s$ functor on a category of classes. The main goal is to establish a soundness theorem as in [18], i.e. we will show that algebraic universes validate the axioms of **IZF**. Moreover, we will analyze how statements formulated in the internal logic of $\mathcal{E}$ can be translated into statements of the language of set theory. We like to think about this as squeezing set theoretic structure into our universe. Finally, we will study a condition ensuring the existence of an initial algebra for the $\mathcal{P}_s$ functor.

Chapter 5 The main goal of the fifth chapter is to show how Cohen's result fits into the context of algebraic set theory. To this end we will follow the presentation of Tierney's proof along the lines of Chapter 6 in [20]. We will translate the Tierney's result in the algebraic set theory framework using the squeezing results that we established earlier.

Chapter 2

Categories of Classes

A category of classes intends to capture the interrelation between classes and sets. It consists of a base category $\mathcal{E}$ together with a distinguished class $\mathcal{S}$ of arrows of $\mathcal{E}$ that are called small maps. The class of small maps determines particular objects of the category $\mathcal{E}$ to be small objects. These small objects in a category of classes intend to behave as sets as opposed to proper classes. The notion of a category equipped with a class of small maps was first introduced by Joyal and Moerdijk [18] and originated in earlier work of the authors.

In this chapter we will introduce the reader to the notion of a category of classes. In the first section we will define the notion of a Heyting pretopos that forms the base category for our further considerations. We will examine elementary properties of these categories, in particular, we will describe how to interpret logic inside a Heyting pretopos. In the next section, we will define the axioms for the class of small maps leading to the definition of a category of classes. In a last section, we will examine the notion of a category of classes more closely. Most of the results presented in this chapter already occur in the original work [18]. Additional results that we present can be found in [28], [30] and [26].

2.1 Heyting pretoposes

The aim of this section is to introduce the reader to the notion of a Heyting pretopos. We will list some basic properties of such categories, in particular, we will show that show that the subobject lattices of objects in a Heyting pretopos form Heyting algebras. This property enables one to interpret logic inside a category. We will explain the details in the following section.

For an object X of $\mathcal{E}$, we denote by $\mathcal{E}/X$, the *slice category of* $\mathcal{E}$ *over* X. The category $\mathcal{E}/X$ has as objects all arrows of $\mathcal{E}$ with codomain X and for $f : A \to X$ and $g : B \to X$ in $\mathrm{Ob}(\mathcal{E}/X)$ every $h : A \to B$ of $\mathcal{E}$ is an arrow from f to g if and only if $f = g \circ h$. If the category $\mathcal{E}$ has pullbacks, every arrow $f : X \to Y$ of $\mathrm{Ar}(\mathcal{E})$ induces a functor

$$f^* : \mathcal{E}/Y \longrightarrow \mathcal{E}/X$$

that acts on objects by pullback along f and on arrows in the obvious way. Observe that f^* always has a left adjoint $\Sigma_f : \mathcal{E}/X \to \mathcal{E}/Y$ given by composition with f.

For every object X of $\mathcal{E}$ we will denote by $\mathrm{Sub}(X)$ *the category of subobjects over* X. Its object are equivalence classes of monos with codomain X where two monos $m : A \to X$ and $n : B \to X$ are equivalent if there are $g : (m : A \to X) \to (n : B \to X)$ and $h : (n : B \to X) \to (m : A \to X)$ in $\mathcal{E}/X$. Arrows in $\mathrm{Sub}(X)$ are inherited from $\mathcal{E}/X$. Observe, that if $g : (m : A \to X) \to (n : B \to X)$ is an arrow in $\mathcal{E}/X$, then g is necessarily a mono in $\mathcal{E}$ and unique as an arrow from m to n. If there is an arrow between A and B in $\mathrm{Sub}(X)$, we will usually write $A \leq B$. The relation $\leq$ makes $\mathrm{Sub}(X)$ into a partial order. Note however, that we will often abuse notation by writing $m : A \to X$ in $\mathrm{Sub}(X)$, when we mean the subobject given by the equivalence class of m. If the category $\mathcal{E}$ has pullbacks then for $f : X \to Y$ the pullback functor from above induces a functor

$$f^{-1} : \mathrm{Sub}(Y) \longrightarrow \mathrm{Sub}(X),$$

since monos are preserved under pullback. If no confusion is to be expected, we will often denote f^{-1} as f^* and call it the pullback functor. Note that f^{-1} does not necessarily have a left adjoint. However, in the case where f itself is a mono, the left adjoint $\exists_f : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ of f^{-1} is induced by composition with f.

Definition 2.1.1. A category $\mathcal{E}$ is called a *pretopos* if

(1) it has all finite limits,

(2) it has finite sums which are stable under pullback and disjoint. The latter means that for every A, B in $\mathcal{E}$, the coproduct inclusions $m_A : A \to A \coprod B$, $m_B : B \to A \coprod B$ are monos and the diagram

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & B \arrow[d] \\ A \arrow[r] & A \coprod B \end{tikzcd}$$

is a pullback.

(3) Equivalence relations have quotients that are stable.

Let us explain the notions used in the definition of a pretopos. First, we will explain the notion of an equivalence relation. A mono $m = \langle m_1, m_2 \rangle : R \to A \times A$ is called an *equivalence relation* if it is

- *reflexive*, i.e. there is $r : A \to R$ such that $m_1 r = m_2 r = \mathrm{id}_A$,
- *symmetric*, i.e. there is $s : R \to R$ such that $m_1 s = m_2$ and $m_2 s = m_1$ and it is
- *transitive*, i.e. if P is the pullback in

$$\begin{tikzcd} P \arrow[r, "q"] \arrow[d, "p"'] & R \arrow[d, "m_1"] \\ R \arrow[r, "m_2"'] & A \end{tikzcd}$$

then there is $t : P \to R$ such that $m_1 t = m_1 p$ and $m_2 t = m_2 q$.

An equivalence relation $\langle m_1, m_2 \rangle : R \to A \times A$ has a *quotient* if there is a map $f : A \to B$ fitting in a diagram of the form

$$\begin{tikzcd} R \arrow[r, shift left, "m_1"] \arrow[r, shift right, "m_2"'] & A \arrow[r, "f"] & B \end{tikzcd}$$

that is both, a pullback and a coequalizer. Next, we will explain what stability under pullback in (2) and (3) of the Definition 2.1.1 means. Assume that $\mathcal{E}$ is a pretopos, X an object of $\mathcal{E}$ and let $h : A \to X$ and $g : B \to X$ be elements of the slice category $\mathcal{E}/X$. It is easy to see that the coproduct $A \coprod B$ of A and B in $\mathcal{E}$ equipped with the universal map $h \coprod g : A \coprod B \to X$ gives rise to the coproduct of h and g in the slice category $\mathcal{E}/X$. Now by stability under pullback, we mean that for every $f : Y \to X$, the pullback functor $f^* : \mathcal{E}/X \to \mathcal{E}/Y$ preserves this structure, i.e. $f^*(h \coprod g : A \coprod B \to X)$ gives rise to the coproduct of $f^*(h : A \to X)$ and $f^*(g : B \to X)$ in $\mathcal{E}/Y$. Similarly, for quotients of equivalence relations.

In fact, it is not difficult to see that if $\mathcal{E}$ is a pretopos, every slice category $\mathcal{E}/X$ has finite limits, disjoint sums, effective equivalence relations and quotients of these. These are again automatically stable under pullback, because every slice category of a slice category is already a slice category of $\mathcal{E}$, see [17, p. 8]. Therefore stability of the under pullback follows in the slice category follows from the corresponding fact in $\mathcal{E}$. We obtain:

Corollary 2.1.2. *If $\mathcal{E}$ is a pretopos, then for every X in $\mathcal{E}$ also the slice category $\mathcal{E}/X$ is a pretopos.*

Next, we are going to list some properties of pretoposes.

Proposition 2.1.3. *Let $\mathcal{E}$ be a pretopos. Then the following properties hold.*

(1) *Every arrow $f : A \to B$ can be factored into a cover, see [17, p. 19] for the definition, followed by a mono. Moreover, the factorization is unique up to unique isomorphism.*

(2) *$\mathcal{E}$ has images, i.e. for every arrow $f : A \to B$ there is a smallest subobject through which f factors. Moreover, images are stable under pullback.*

(3) *For every morphism $f : Y \to X$, the pullback functor $f^{-1} : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ has a left adjoint $\exists_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$. Moreover, the adjunction satisfies the Beck-Chevalley condition, i.e. for every pullback diagram*

$$\begin{tikzcd} W \arrow[r, "f"] \arrow[d, "j"'] & X \arrow[d, "h"] \\ Y \arrow[r, "g"'] & Z \end{tikzcd}$$

and A in $\mathrm{Sub}(X)$, we have that $g^{-1}(\exists_h(A)) = \exists_j(f^{-1}(A))$ in $\mathrm{Sub}(Y)$.

(4) *The initial object 0 of $\mathcal{E}$ is strict, i.e. every arrow $f : A \to 0$ is an isomorphism.*

(5) *Every mono and every epi in $\mathcal{E}$ is regular. Moreover, $\mathcal{E}$ is balanced, i.e. every arrow that is mono and epi is an isomorphism.*

Proof. Although all the verification of the properties listed are fairly standard, we would still like to sketch some of them.

(1) A worked out proof of this statement can be found in [24]. The construction is as follows. Let $f : A \to B$ be an arrow in $\mathcal{E}$. It is easy to see that the kernel pair (p_1, p_2) of f, i.e. the maps fitting in the pullback

$$\begin{tikzcd} Z \arrow[r, "p_1"] \arrow[d, "p_2"'] & A \arrow[d, "f"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

defines an equivalence relation on A. By (3) of Definition 2.1.1 let $e : A \to Q$ be the coequalizer of this relation. Then e is clearly a regular epi and therefore a cover as regular epis are always covers. Since $f \circ p_1 = f \circ p_2$ there is a unique map $m : Q \to B$ such that $f = m \circ e$. Now one can show that m is a mono which gives the desired factorization.

(2) For an arrow $f : A \to B$ let

$$A \xrightarrow{e} Q \xrightarrow{m} B$$

be the unique factorization as above, where e is the coequalizer of the kernel pair (p_1, p_2) of f. It is easy to see that $m : Q \to B$ is the smallest subobject through which f factors and therefore provides the image of f. Stability follows from the fact that quotients of equivalence relations are stable.

(3) Let $f : Y \to X$ be an arrow in $\mathcal{E}$. We define

$$\exists_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X).$$

as follows. Given a subobject $m_C : C \to Y$ of Y let $\exists_f(m_C : C \to Y)$ be the image of $f \circ m_C$ as in

$$\begin{tikzcd} C \arrow[r, tail, "m_C"] \arrow[dr, two heads, "e"'] & Y \arrow[r, "f"] & X \\ & \exists_f(C) \arrow[ur, tail, "m"'] & \end{tikzcd} .$$

From the fact that such a factorization is unique, it follows that $\exists_f$ is well-defined. It is also easy to see that this assignment is functorial, i.e. order preserving. For let $m_d : D \to Y$ and $m_c : C \to Y$ in $\mathrm{Sub}(Y)$ with

$D \leq C$ witnessed by $k : D \to C$, i.e. we have $m_c \circ k = m_d$. By the latter it is clear that $f \circ m_d$ factors through $\exists_f(C)$ but this implies that $\exists_f(D) \leq \exists_f(C)$ in $\mathrm{Sub}(X)$.

We show that $\exists_f$ is a left adjoint to f^{-1}, i.e. for all $m_C : C \to Y \in \mathrm{Sub}(Y)$ and $m_A : A \to X \in \mathrm{Sub}(X)$ we have

$$\exists_f(C) \leq A \quad \Leftrightarrow \quad C \leq f^{-1}(A).$$

For the direction from left to right, assume that there is a map $h : \exists_f(C) \to A \in \mathrm{Sub}(X)$. Then $f \circ m_C = m \circ e = m_A \circ h \circ e$, where $m \circ e$ is the factorization of $f \circ m_C$. So by the universal property of the pullback in

$$\begin{tikzcd} f^{-1}(A) \arrow[r, "p_2"] \arrow[d, "m'"'] & A \arrow[d, "m_A"] \\ Y \arrow[r, "f"'] & X \end{tikzcd}$$

there is a unique map $g : C \to f^{-1}(A)$ with $m' \circ g = m_C$. This witnesses that $C \leq f^{-1}(A) \in \mathrm{Sub}(Y)$. Conversely, assume there is a map $k : C \to f^{-1}(A) \in \mathrm{Sub}(Y)$. But this implies that $f \circ m_C$ factors through $m_A : A \to X$. Since $\exists_f(C)$ is the smallest subobject of X with this property, it follows that $\exists_f(C) \leq A$ in $\mathrm{Sub}(X)$. This finishes the proof that $\exists_f$ is left adjoint to f^{-1}. That the adjunction satisfies the Beck-Chevalley condition follows from the fact that images are preserved under pullback.

(4) This is proven in [17, p. 31] for coherent categories. The proof uses that for every X, $\mathrm{Sub}(X)$ has a smallest object that it stable under pullback. Since $\mathcal{E}$ has finite sums, it follows that $\mathcal{E}$ has an initial object 0. For every X the image of the unique map $0 \to X$ is easily seen to be the smallest object of $\mathrm{Sub}(X)$. Stability follows from stability of finite sums and images.

(5) See [17, p. 38].

□

Definition 2.1.4. A pretopos $\mathcal{E}$ is a *Heyting pretopos* if for every $f : Y \to X$ of $\mathcal{E}$ the pullback functor $f^{-1} : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ has a right adjoint $\forall_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$.

The next proposition tells us that in a Heyting pretopos, $\mathrm{Sub}(X)$ carries the structure of a Heyting algebra. This already indicates the logical character of such categories. In the next section, we will explain in detail how a Heyting pretopos provides a semantics that is sound with respect to first order intuitionistic logic.

Proposition 2.1.5. *Let $\mathcal{E}$ be a pretopos. Then every subobject poset forms a distributive lattice and the lattice structure is preserved under pullback. If $\mathcal{E}$ is a Heyting pretopos then these lattices are Heyting algebras.*

Proof. Let X be an object of $\mathcal{E}$. Given subobjects $m_A : A \to X$ and $m_B : B \to X$ of X, define the meet of A and B as the pullback in the following diagram.

$$\begin{tikzcd} A \wedge B \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & B \arrow[d, "m_B"] \\ A \arrow[r, "m_A"'] & X \end{tikzcd}$$

Since the pullback of a mono is a mono, π_1 is a mono, and therefore $m_A \circ \pi_1 : A \wedge B \to X$ witnesses that $A \wedge B$ is indeed a subobject of X. Now the universal property of the pullback corresponds precisely to the fact that $A \wedge B$ is the greatest lower bound of A and B in $\mathrm{Sub}(X)$. We already remarked that such pullbacks are preserved by the pullback functors. The meet of A and B in $\mathrm{Sub}(X)$ is constructed in two steps. First, we take the sum of A and B and obtain a canonical map from $s : A \coprod B \to X$ by the universal property of finite sums. The map s is not necessarily a mono. We define $A \vee B$ as the image of s. Then it is not difficult to see that $A \vee B$ defines the least upper bound of A and B. Stability under pullback follows from the fact that in pretopoi finite sums and images are preserved under pullback. In Proposition 2.1.3, (3), we argued that the initial object 0 of $\mathcal{E}$ is strict. An easily verified consequence of strictness is that the unique arrow $0 \to X$ is a mono. It is clear that this defines the least object of $\mathrm{Sub}(X)$. Moreover, the identity arrow $\mathrm{id}_X : X \to X$ serves as the greatest element in the lattice. The distributivity of the lattice is proved in [17, p. 32].

Now assume that $\mathcal{E}$ is a Heyting pretopos. It remains to show that we also have a Heyting implication that is preserved under pullback. So let $m : A \to X$ and $n : B \to X$ be elements of $\mathrm{Sub}(X)$. Since $\mathcal{E}$ is a Heyting pretopos, the pullback functor $m^{-1} : \mathrm{Sub}(X) \to \mathrm{Sub}(A)$ has a right adjoint $\forall_m : \mathrm{Sub}(A) \to \mathrm{Sub}(X)$. Define

$$A \Rightarrow B := \forall_m(m^{-1}(B)).$$

In order to show that $A \Rightarrow B$ satisfies the axiom for Heyting implication, we need to show that for every C in $\mathrm{Sub}(X)$, $A \wedge C \leq B$ if and only if $C \leq A \Rightarrow B$. Observe, that since $m : A \to X$ is a mono, the left adjoint $\exists_m : \mathrm{Sub}(A) \to \mathrm{Sub}(X)$ is simply given by postcomposition with m. This shows that for every C in $\mathrm{Sub}(X)$, $A \wedge C = \exists_m(m^{-1}(C))$. We then have for some $C \in \mathrm{Sub}(X)$:

$$\begin{array}{llll} & A \wedge C \leq B & & \\ \Leftrightarrow & \exists_m(m^{-1}(C)) \quad \leq & B & [\text{in } \mathrm{Sub}(X)] \\ \Leftrightarrow & m^{-1}(C) \quad \leq & m^{-1}(B) & [\text{in } \mathrm{Sub}(A), \text{ since } \exists_m \vdash m^{-1}\ ] \\ \Leftrightarrow & C \quad \leq & \forall_m(m^{-1}(B)) & [\text{in } \mathrm{Sub}(X)] \\ \Leftrightarrow & C \leq A \Rightarrow B. & & \end{array}$$

This finishes the proof that $A \Rightarrow B$ satisfies the properties of Heyting implication. We already argued that left adjoints of the pullback functors satisfy the Beck-Chevalley condition. From this it easily follows that also the right adjoint satisfy this property, i.e. commutes with the pullback functor, as the following calculation shows. Let

$$\begin{array}{ccc} W & \xrightarrow{p} & X \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[g]{} & Z \end{array}$$

be a pullback. Assume B in $\mathrm{Sub}(X)$. Then

$$\begin{array}{lll} & \forall_p q^{-1}(B) \leq \forall_p q^{-1}(B) & [\text{in } \mathrm{Sub}(X)] \\ \Rightarrow & p^{-1}\forall_p q^{-1}(B) \leq q^{-1}(B) & [\text{since } p^{-1} \vdash \forall_p] \\ \Rightarrow & \exists_q p^{-1}\forall_p q^{-1}(B) \leq B & [\text{since } \exists_q \vdash q^{-1}] \\ \Rightarrow & g^{-1}\exists_f \forall_p q^{-1}(B) \leq B & [\text{since } \exists_q \vdash q^{-1} \text{ satisfies B.-Ch.}] \\ \Rightarrow & \forall_p q^{-1}(B) \leq f^{-1}\forall_g(B) & [\text{adjunction properties.}] \end{array}$$

The other direction, that $f^{-1}\forall_g(B) \leq \forall_p q^{-1}(B)$ follows similarly, finishing the proof that the adjunction satisfies the Beck-Chevalley condition. It is immediate to infer that Heyting implication is preserved under pullback. □

Definition 2.1.6. A pretopos is called a *Boolean pretopos* if all subobject lattices are Boolean algebras.

From the definition of a Boolean pretopos it is not immediately clear that Boolean pretoposes form a special case of Heyting pretoposes. This fact is established in the next proposition.

Proposition 2.1.7. *Every Boolean pretopos is a Heyting pretopos.*

Proof. Let $\mathcal{E}$ be a Boolean pretopos. Since $\mathcal{E}$ is in particular a pretopos, it remains to prove that the pullback functors have right adjoints. By Proposition 2.1.5 it follows that every subobject poset is a distributive lattice and the lattice structure is preserved under pullback. Observe that we can infer that also the complement operation is preserved under pullbacks. To this end let $f : Y \to X$ and arrow in $\mathcal{E}$ and $A \in \mathrm{Sub}(X)$. Then

$$Y = f^*(X) = f^*(A \vee A^C) = f^*(A) \vee f^*(A^C). \tag{2.1}$$

Since sums are disjoint, it follow that $f^*(A^C)$ is the complement of $f^*(A)$ in $\mathrm{Sub}(Y)$, i.e. $f^*(A^C) = f^*(A)^C$. In Proposition 2.1.5 we also showed that f^* has a left adjoint $\exists_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$. We make use of this in order to define a functor

$$\forall_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X) \text{ by } B \mapsto \exists_f(B^C)^C.$$

Since complements are unique, the assignment is well-defined and one can easily check that it is functorial. Then for all $A \in \mathrm{Sub}(X)$, $B \in \mathrm{Sub}(Y)$ we have

$$\begin{array}{lll} & f^*(A) \leq B & (\text{in } \mathrm{Sub}(Y)) \\ \Leftrightarrow & B^C \leq f^*(A)^C = f^*(A^C) & [\text{properties of the complement and (1)}] \\ \Leftrightarrow & \exists_f(B^C) \leq A^C & [\text{since } \exists_f \dashv f^*] \\ \Leftrightarrow & A \leq \left(\exists_f(B^C)\right)^C = \forall_f(B) & [\text{complement properties}]. \end{array}$$

This shows that $\forall_f$ is indeed the right adjoint of f. □

2.1.1 Interpreting First Order Intuitionistic Logic inside a Heyting Pretopos

In this section, we will show how to interpret first order languages in a Heyting pretopos giving rise to a sound semantics with respect to first order intuitionistic logic. Moreover, we will see that Boolean pretoposes give rise to sound semantics with respect to classical logic. Formulas of a first-order (sorted) language will always be interpreted in a context, i.e. in a string $\bar{x} = x_1 : X_1, \dots x_n : X_n$ of sorted variables. The sorts of the language will be interpreted as objects of the category. Assuming φ to be a formula all whose free variables are among $x_1 : X_1, \dots x_n : X_n$, then the interpretation of φ in the context $x_1 : X_1, \dots x_n : X_n$ is defined to be a subobject of the product of the interpretation of the sorts occurring in the context.

The results presented in this section summarize the treatment of first order categorical logic in Johnstone's book [17]. For a more thorough introduction to categorical logic the reader is referred to [21] or [25].

We will now make the ideas described above more explicit. First, we define how to interpret languages in a Heyting pretopos.

Definition 2.1.8. Let $\mathcal{E}$ be a Heyting pretopos and let $\mathcal{L}$ be a sorted first-order language. An *interpretation of* $\mathcal{L}$ *in* $\mathcal{E}$ is provided if

- To every sort X of $\mathcal{L}$ is assigned an object $M(X)$ of $\mathcal{E}$. Note that we will often abuse notation and denote the name of the sort X and its interpreting object $M(X)$ with the same letter X.

- Every relation symbol $R : X_1, \dots, X_n$ is assigned to a subobject

$$[\![R]\!] \rightarrowtail M(X_1) \times \cdots \times M(X_n)$$

- Every function symbol $f : X_1, \dots, X_n \to Y$ of $\mathcal{L}$ is assigned to an arrow

$$[\![f]\!] : M(X_1) \times \cdots \times M(X_n) \to M(Y) \text{ in } \mathcal{E}\,.$$

Having an interpretation of the language $\mathcal{L}$ in $\mathcal{E}$, we get an interpretation of terms of the language $\mathcal{L}$. A term t will always be interpreted in a context $\bar{x}$ where all free variables of t are among the ones in $\bar{x}$. If this condition is satisfied, we say that the particular *context is suitable* for the term.

Definition 2.1.9. For a term $\bar{x}.t$ in context, i.e. a term t in a suitable context $\bar{x}$, we define its interpretation recursively:

- If $t = x_i : X_i$ a variable, then

$$[\![\bar{x}.x_i]\!] : X_1 \times \cdots \times X_n \to X_i$$

  is the i-th projection.

- If $t = ft_1 \dots t_m$, $f : Y_1, \dots, Y_n \to Z$ and $t_i : Y_i$, then the interpretation $[\![\bar{x}.t]\!]$ is given by the composite

$$X_1 \times \cdots \times X_n \xrightarrow{\langle [\![\bar{x}.t_1]\!], \dots, [\![\bar{x}.t_m]\!] \rangle} Y_1 \times \cdots \times Y_m \xrightarrow{[\![f]\!]} Z\,.$$

By a straightforward induction, one can prove that this definition enjoys the *substitution property.*

Lemma 2.1.10. *Let $\bar{x}$ be a suitable context for a term $t : Z$. Let $s_i : X_i$ be a string a terms of the same length as $\bar{x}$. Let $\bar{y}$ be a suitable context for all the terms in this string. Then the interpretation $[\![\bar{y}.t[\bar{s}/\bar{x}]]\!]$ is given as the composite*

$$Y_1 \times \cdots \times Y_m \xrightarrow{\langle [\![\bar{y}.s_1]\!], \ldots, [\![\bar{y}.s_n]\!] \rangle} X_1 \times \cdots \times X_n \xrightarrow{[\![\bar{x}.t]\!]} Z \; .$$

Observe that by using a trivial substitution, the above Lemma can also be applied if only some of the variables in t are substituted.

We will now sketch how to interpret first order formulas in our category. Note that in our context equality will always be part of the language. As in the case of terms, also formulas are interpreted in a context of variables. Again, we call a *context $\bar{x}$ suitable for a formula φ* if all free variables of φ are among the ones in $\bar{x}$. As already mentioned in the introduction of this section, the interpretation of a formula φ in context $\bar{x}$ will be a subobject of $X_1 \times \cdots \times X_n$, where $X_1, \ldots, X_n$ are the interpretations of the sorts of the variables in $\bar{x}$. Note that if variables x_i and x_j have the same sort then the interpretation of this sort occurs multiple times in the product. One can think of the subobject $[\![\bar{x}.\varphi]\!]$ as containing exactly those elements of $X_1 \times \cdots \times X_n$ that satisfy the formula φ.

Definition 2.1.11. We define the interpretation of a formula φ in suitable context $\bar{x}$ by induction over φ.

Atomic formulas:

If $\bar{x}$ is a suitable context for terms $t : Z$ and $s : Z$, the interpretation $[\![\bar{x}.t = s]\!]$ is the equalizer in

$$\begin{tikzcd} {[\![\bar{x}.t = s]\!]} \arrow[r, tail] & X_1 \times \cdots \times X_n \arrow[r, shift left, "{[\![\bar{x}.t]\!]}"] \arrow[r, shift right, "{[\![\bar{x}.s]\!]}"'] & Z \end{tikzcd} \; .$$

If $R : Y_1, \ldots, Y_m$ is a relation symbol, $t_i : Y_i$ terms, and $\bar{x}$ a suitable context for all these terms, then the interpretation of the formula $Rt_1, \ldots t_m$ is the pullback in

$$\begin{tikzcd} {[\![\bar{x}.Rt_1, \ldots, t_m]\!]} \arrow[r] \arrow[d, tail] & {[\![R]\!]} \arrow[d, tail] \\ X_1 \times \cdots \times X_n \arrow[r, "{\langle [\![\bar{x}.t_1]\!], \ldots, [\![\bar{x}.t_m]\!] \rangle}"] & Y_1 \times \cdots \times Y_m \end{tikzcd} \; .$$

Boolean connectives:

The algebraic structure on $\mathrm{Sub}(X_1 \times \cdots \times X_n)$ according to Proposition 2.1.5 allows us to interpret Boolean connections of formulas. So if $\bar{x}$ is a suitable context for formulas φ and ψ and we already have interpretations $[\![\bar{x}.\varphi]\!]$, $[\![\bar{x}.\varphi]\!]$ in $\mathrm{Sub}(X_1 \times \cdots \times X_n)$ we define

- $[\![\bar{x}.\bot]\!] := 0$ and $[\![\bar{x}.\top]\!] := 1$, the largest subobject of $\mathrm{Sub}\, X_1 \times \cdots \times X_n$.

- $[\![\bar{x}.\varphi \wedge \psi]\!] := [\![\bar{x}.\varphi]\!] \wedge [\![\bar{x}.\psi]\!]$
- $[\![\bar{x}.\varphi \vee \psi]\!] := [\![\bar{x}.\varphi]\!] \vee [\![\bar{x}.\psi]\!]$
- $[\![\bar{x}.\varphi \Rightarrow \psi]\!] := [\![\bar{x}.\varphi]\!] \Rightarrow [\![\bar{x}.\psi]\!]$
- $[\![\bar{x}.\neg\varphi]\!] := [\![\bar{x}.\varphi]\!] \Rightarrow 0$,

where the operations on the right hand side arise from the Heyting algebra structure on $\mathrm{Sub}(X_1 \times \cdots \times X_n)$.

Quantifiers:

Assume $y\bar{x}$ is a suitable context for the formula φ and we already have an interpretation $[\![y\bar{x}.\varphi]\!]$ in $\mathrm{Sub}(Y \times X_1 \times \cdots \times X_n)$. Then $\bar{x}$ is a suitable context for the formulas $\exists y\varphi$ and $\forall y\varphi$. Let

$$\pi_Y : Y \times X_1 \times \cdots \times X_n \longrightarrow X_1 \times \cdots \times X_n$$

denote the obvious projection map. The formulas $\exists y\varphi$ and $\forall y\varphi$ are interpreted as

- $[\![\bar{x}.\exists y\varphi]\!] := \exists_{\pi_Y}([\![y\bar{x}.\varphi]\!])$ and
- $[\![\bar{x}.\forall y\varphi]\!] := \forall_{\pi_Y}([\![y\bar{x}.\varphi]\!])$.

Here $\exists_{\pi_Y}$ and $\forall_{\pi_Y}$ are the left and right adjoints of the pullback functor associated to π_Y.

In Proposition 2.1.5 we did not only show that the subobjects lattices in $\mathcal{E}$ have a Heyting algebra structure, we also showed that this structure should be preserved under pullbacks. Moreover we showed that the left and right adjoints of the pullback functor satisfy the Beck-Chevalley condition, i.e. commute with the pullback operations. These properties are exactly the ones we need to prove the *substitution property for formulas.*

Lemma 2.1.12. *Let $\bar{x}$ be a suitable context for φ and let $t_i : X_i$ a string of terms of the same length as $\bar{x}$. Moreover let $\bar{y}$ be a suitable context for all the term in the string $\bar{t}$. Then $[\![\bar{y}.\varphi(t_i/x_i)]\!]$ is the pullback in*

$$\begin{tikzcd}
{[\![\bar{y}.\varphi(t_i/x_i)]\!]} \arrow[r] \arrow[d, tail] & {[\![\bar{x}.\varphi]\!]} \arrow[d, tail] \\
Y_1 \times \cdots \times Y_m \arrow[r, "{\langle [\![\bar{y}.t_1]\!], \ldots, [\![\bar{y}.t_n]\!] \rangle}"] & X_1 \times \cdots \times X_n
\end{tikzcd}.$$

The above lemma also sheds light on how formulas behave by enlarging the context. Assume φ is a formula with free variables $\bar{x} = x_1, \ldots, x_n$. Let $\bar{y}$ be a string of variables such that all variables in $\bar{x}$ occur in the string $\bar{y}$. Then $\bar{y}$ is also a suitable context for φ. Applying the trivial substitution, i.e. substitute variables x_i for x_i, Lemma 2.1.12 gives that the interpretation of φ in context $\bar{y}$ is given by the pullback in

$$\begin{tikzcd}
{[\![\bar{y}.\varphi]\!]} \arrow[r] \arrow[d, tail] & {[\![\bar{x}.\varphi]\!]} \arrow[d, tail] \\
Y_1 \times \cdots \times Y_m \arrow[r, "{\langle [\![\bar{y}.x_1]\!], \ldots, [\![\bar{y}.x_n]\!] \rangle}"] & X_1 \times \cdots \times X_n
\end{tikzcd}.$$

In light of the semantics that we will define below, this shows soundness of the weakening rule. If a formula φ does not contain any free variables, then we call φ a *sentence*. In this case, the empty context is suitable for φ and $[\![\varphi]\!]$ is interpreted as an element in $\mathrm{Sub}(1)$, where 1 is the terminal object of the category $\mathcal{E}$.

Finally, we will provide a definition of truth of a sequent in a Heyting pretopos.

Definition 2.1.13. Let $\mathcal{E}$ be a Heyting pretopos, $\mathcal{L}$ a sorted language and assume we have an interpretation of $\mathcal{L}$ in $\mathcal{E}$. Let φ and ψ be $\mathcal{L}$-formulas in context $\bar{x}$ that is suitable for φ and ψ.

- We say that a *sequent $\varphi \vdash_{\bar{x}} \psi$ in context $\bar{x}$ is satisfied in $\mathcal{E}$* if and only if
$$[\![\bar{x}.\varphi]\!] \leq [\![\bar{x}.\psi]\!] \text{ in } \mathrm{Sub}(X_1 \times \cdots \times X_n).$$
In this case we write: $\mathcal{E} \models_{\bar{x}} \varphi \vdash \psi$.
- Note that if φ on the right hand side is $\top$ then the sequent $\top \vdash_{\bar{x}} \psi$ is satisfied in $\mathcal{E}$ if
$$[\![\bar{x}.\psi]\!] \cong X_1 \times \cdots \times X_n \text{ in } \mathrm{Sub}(X_1 \times \cdots \times X_n).$$
In this case we write: $\mathcal{E} \models_{\bar{x}} \psi$.
- In particular, if φ is a sentence, then $\mathcal{E} \models \varphi$ iff $[\![\varphi]\!] \cong 1$ in $\mathrm{Sub}(1)$.
- If $\mathbb{T}$ is a first order theory, i.e. a set of sequences of the form $\varphi \vdash_{\bar{x}} \psi$. We say that $\mathcal{E}$ is a *model for* $\mathbb{T}$ if all sequences of $\mathbb{T}$ are satisfied in $\mathcal{E}$.

Note that we will often omit the free variables $\bar{x}$ in the notation $\mathcal{E} \models_{\bar{x}} \varphi$ when no confusion is to be expected. The above definition provides a semantics that is sound with respect to first order intuitionistic logic, see for example [17, p. 830] for a suitable calculus.

Proposition 2.1.14 (Soundness)**.** *Let $\mathbb{T}$ be a first-order theory. Let $\mathcal{E}$ be a Heyting pretopos that is a model for $\mathbb{T}$. Let $\varphi \vdash \psi$ be a sequent. Then*
$$(\mathbb{T} \vdash_i \varphi \vdash \psi) \quad \Longrightarrow \quad (\mathcal{E} \models \varphi \vdash \psi)\,.$$

Proof. A proof can be found in [17, p. 823]. Note, however, that we have actually already done most of the work. For example, to see that the substitution rule is sound use Lemma 2.1.12 and the fact that pullback functors are order preserving, the soundness for the cut-rule follows by transitivity of the partial order on the subobject lattices. The rules for Boolean connectives are sound by the Heyting algebra structure on the subobject lattices. Finally, the rules for the quantifiers are:
$$\frac{\varphi \vdash_{\bar{x},y} \psi}{\overline{\exists y \varphi \vdash_{\bar{x}} \psi}} \qquad \frac{\psi \vdash_{\bar{x},y} \varphi}{\overline{\psi \vdash_{\bar{x}} \forall y \varphi}}$$
The notation implies that $\bar{x}$ is a suitable context for the formula ψ. Our considerations above Definition 2.1.13 revealed that $[\![\bar{x}y.\psi]\!] = \pi_y^{-1}([\![\bar{x}.\psi]\!])$, where $\pi_Y : Y \times X_1 \times \cdots \times X_n \to X_1 \times \cdots \times X_n$ is the projection map as in Definition 2.1.8. Now it is clear that the quantification rules express exactly the properties of right and left adjoints of the pullback functor associated to π_Y.[1] $\square$

[1]The observation that quantification can be modeled categorically by right and left adjoints

2.1.2 The Internal Language of a Heyting pretopos

Above, we saw how to interpret first order languages in a category. Every category comes in fact with its own internal language that allows us to reason about the category itself in a formal way.

Definition 2.1.15. Let $\mathcal{E}$ be a category with finite products. The *internal language of* $\mathcal{E}$ is given by:

- The sorts of $\mathcal{L}(\mathcal{E})$ are given by the objects of the category.
- Every arrow $f : X_1 \times \cdots \times X_n \to Y$ in $\mathcal{E}$ determines a function symbol $f : X_1, \ldots, X_n \to Y$ of $\mathcal{L}(\mathcal{E})$.
- And every mono $R \to X_1 \times \cdots \times X_n$ determines a relation symbol $R : X_1, \ldots, X_n$.

The internal language $\mathcal{L}(\mathcal{E})$ has a canonical interpretation in $\mathcal{E}$ by interpreting the sorts, function and predicate symbols by the objects, functions and subobjects that gave rise to them. Since $\mathcal{E}$ is a Heyting pretopos, the soundness theorem (Proposition 2.1.14) tells us that sequences in the internal language are closed under the rules of inference of intuitionistic logic. This provides us a way to prove statements about the category $\mathcal{E}$ in a formal language. In fact, this way of proving properties of $\mathcal{E}$ can replace the diagrammatic reasoning in which arguments are often hard to follow [25, p. 3]. We will see many instances of this reasoning later.

For now, let us see some elementary examples how the internal language can be used on the one hand to characterize subobjects and on the other hand to prove properties about $\mathcal{E}$. The following examples are taken from [21].

Characterizing subobjects:

(1) Assume $f : A \to X$ is a mono. Then f is a subobject of X. One can easily see that this subobject is determined by the interpretation of the formula $[\![x.\exists a(f(a) = x)]\!]$

(2) Given arrows $f : Y \to X$ and $g : Z \to X$, the formula $f(y) = g(z)$ determines a subobject $[\![yz.f(y) = g(z)]\!]$ of $Y \times Z$. Now it is easy to see that this subobject corresponds to the one given by the mono $\langle \pi_Y, \pi_Z \rangle : Y \times_X Z \to Y \times Z$, where π_Y and π_Z are coming from the pullback in

$$\begin{array}{ccc} Y \times_X Z & \xrightarrow{\pi_Z} & Z \\ {\scriptstyle \pi_Y}\downarrow & & \downarrow{\scriptstyle g} \\ Y & \xrightarrow[f]{} & X \end{array} .$$

Expressing properties of $\mathcal{E}$:

(1) $f : X \longrightarrow X$ is the identity morphism if and only if $\mathcal{E} \models_x f(x) = x$.

---

of a weakening functor was first made by Lawvere and was a milestone in the development of categorical logic.

(2) $f : X \longrightarrow Y$ is monic if and only if $\mathcal{E} \models_{xx'} (f(x) = f(x') \vdash x = x')$

Let us sketch the proof of (2). For the direction from left to right assume that $f : X \to Y$ is monic. As already remarked above, the subobject $[\![xx'.f(x) = f(x')]\!]$, is given by $\langle p_1, p_2\rangle : [\![xx'.f(x) = f(x')]\!] \to X \times X$, where p_1 and p_2 are coming from the pullback

$$\begin{tikzcd} {[\![xx'.f(x) = f(x')]\!]} \arrow[r, "p_1"] \arrow[d, "p_2"'] & X \arrow[d, "f"] \\ X \arrow[r, "f"'] & X \end{tikzcd}.$$

The subobject $[\![xx'.x = x']\!]$ is given by the diagonal $\Delta : X \to X \times X$. Since f is a mono and $fp_1 = fp_2$ it follows that $p_1 = p_2$. But this shows that $\langle p_1, p_2\rangle$ factors through the diagonal and therefore $[\![xx'.f(x) = f(x')]\!] \leq [\![xx'.x = x']\!]$ in $\mathrm{Sub}(X \times X)$, which shows that the right hand side of (2) is satisfied. For the direction from right to left, assume that $[\![xx'.f(x) = f(x')]\!] \leq [\![xx'.x = x']\!]$ in $\mathrm{Sub}(X \times X)$. This implies that $p_1 = p_2$, where p_1 and p_2 are as in the pullback above. In order to show that f is mono, let $g, h : W \to Y$ be arrows such that $fg = fh$. By the universal property of the pullback above there is a unique map $k : W \to [\![xx'.f(x) = f(x')]\!]$ with $g = p_1k$ and $h = p_2k$. Since $p_1 = p_2$ it follows that $g = h$. This shows that f is mono.

We finish this section by stating a very useful lemma that tells us that internally definable functions in the category $\mathcal{E}$ indeed occur as arrows in $\mathcal{E}$. For a proof of the lemma the reader is referred to [21, p. 89].

Lemma 2.1.16. *Assume $R \to A \times B$ a mono, is provably a functional relation, i.e.*

$$\mathcal{E} \models_{a.A,b:B} (Rab \wedge Rab') \to b = b' \text{ and}$$
$$\mathcal{E} \models_{a:A} \exists b : BRab$$

then there is a unique arrow $f : A \to B$ such that the graph of f, $\Gamma(f) := [\![a : Ab : B \mid f(a) = b]\!] = R$ in $\mathrm{Sub}(A \times B)$.

Another useful lemma concerns the truth of existential statements.

Lemma 2.1.17. *Assume φ is a formula such that all free variables of φ are among $a : A, \bar{x} : \bar{X}$ and $\bar{y} : \bar{Y}$. Assume $f : \bar{X} \to A$ is an arrow in $\mathcal{E}$. Then if*

$$\mathcal{E} \models \varphi[f(\bar{x})/a] \text{ then } \mathcal{E} \models \exists a : A\varphi.$$

Proof. All relevant interpretations are shown in the diagram below, where the square is a pullback and m is the image of $\langle \pi_{\bar{Y}}, \pi_{\bar{X}}\rangle n$.

$$\begin{tikzcd}
\llbracket \varphi[f(\bar{x})/a] \rrbracket \arrow[r] \arrow[d, "\cong"'] & \llbracket \varphi \rrbracket \arrow[d, tail, "n"'] \arrow[dr, two heads, "e"] & \\
\bar{Y} \times \bar{X} \arrow[r, "{\langle \pi_{\bar{Y}}, \pi_{\bar{X}}, f\pi_{\bar{X}} \rangle}"] \arrow[dr, no head] & \bar{Y} \times \bar{X} \times A \arrow[d, "{\langle \pi_{\bar{Y}}, \pi_{\bar{X}} \rangle}"'] & \llbracket \exists a : A\varphi \rrbracket \arrow[dl, tail, "m"] \\
& \bar{Y} \times \bar{X} &
\end{tikzcd}$$

Since $\mathcal{E} \models \varphi[f(\bar{x})/a]$, the map on the left hand side is iso. It follows that $\llbracket \exists a : A\varphi \rrbracket = \bar{Y} \times \bar{X}$, and we conclude that $\mathcal{E} \models \exists a : A\varphi$. That proves the claim.
□

2.2 Axioms for the Small Maps

In this section, we are going to introduce the notion of small maps on a Heyting pretopos $\mathcal{E}$ with a stable natural numbers object[2]. The small maps provide an abstract notion of being set-sized as opposed to be a proper class. Instead of defining a notion of smallness for objects, we will define smallness of maps. In this way, we automatically obtain the notion of small maps in all slice categories. The reader is invited to think about a small map as being fibrewise small. For ordinary set theory this means that a map $f : X \to Y$ between classes is small if for every $y \in Y$ the preimage $f^{-1}(y)$ is a set. Having a notion of small maps at hand, the small objects are defined to be the ones whose unique map to the terminal object is small. The axiom systems for the small maps the reader may find in the current literature on algebraic set theory vary. Compare for example the choices made in the original approach [18] with the choices made in [28] or [5]. One reason is that different axiomatizations lead to soundness with respect to different set theories. In particular predicative aspects can be considered. This flexibility is in fact one great feature of the categorical framework to set theory. Our choice of the axioms for the small maps is led by the goal to obtain models for the set theory **IZF** and **ZF** in the case that the underlying Heyting pretopos is Boolean. Moreover, we feel that in our presentation the relation to the set theoretic axioms is most evident.

In the following, let $\mathcal{E}$ always be a Heyting pretopos with a stable natural number object.

Definition 2.2.1. A class $\mathcal{S}$ of arrows of $\mathcal{E}$ is called a *class of small maps* if it satisfies the following conditions:

(S1) All monos belong to $\mathcal{S}$.

(S2) The maps in $\mathcal{S}$ are closed under composition.

(S3) If in the diagram below, g is small, e an epi then also f is small.

[2]See [17, p. 108] for the definition of a stable natural numbers object.

$$\begin{tikzcd} X \arrow[rr, "g"] \arrow[dr, "e"'] & & Y \\ & Z \arrow[ur, "f"'] & \end{tikzcd}$$

(S4) The maps in $\mathcal{S}$ are stable under pullback, i.e. if in the pullback

$$\begin{tikzcd} A \arrow[r] \arrow[d, "g"'] & B \arrow[d, "f"] \\ C \arrow[r] & D \end{tikzcd}$$

f is small then also g is small.

(S5) If $f : A \to X$ and $g : B \to X$ are small maps, then also $f + g : A + B \to X$ is small, where $f + g$ arises from the universal property of the coproduct.

As already mentioned, having the notion of small maps at hand we can define small objects as the ones whose unique map to the terminal object is small. The small maps also determine small relations.

Definition 2.2.2.
- An object $A \in \mathrm{Ob}(\mathcal{E})$ is called *small*, if the unique map $A \to 1$ to the terminal object is small.
- A relation $R \subseteq X \times Y$ is called *small in* X or just *small*, if the composite

$$R \rightarrowtail X \times Y \xrightarrow{\pi_2} Y$$

is a small map.

One can think of a relation R on $X \times Y$ as small if for any $y \in Y$ the R predecessors of y build a set. In the context of set theory such a relation is called set-like. Note that the order of the product matters, so a relation R on $X \times Y$ that is small in X is not necessarily small in Y.

A next axiom that a category of classes is supposed to satisfy concerns the representability of small relations. It in particular requires that for every object of the Heyting pretopos there is powerclass object.

(P1) For any object X there is an object $\mathcal{P}_{\mathrm{s}}(X)$ and a small relation $\in_X \subseteq X \times \mathcal{P}_s(X)$ such that for any small relation $R \subseteq X \times Y$ there is a unique map $t : Y \to \mathcal{P}_{\mathrm{s}}(X)$ which fits into a pullback of the form:

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] & \in_X \arrow[d, tail] \\ X \times Y \arrow[r, "\mathrm{id}_X \times t"] & X \times \mathcal{P}_{\mathrm{s}}(X) \end{tikzcd}$$

Intuitively, axiom (P1) states that every small relation $R \subseteq X \times Y$ can be written as $x \in t(y)$, where $t(y) = \{x \mid xRy\}$, i.e. the relation R is representable by the relation $\in_X$. We will later express this precisely in the internal language of the category.

Using the internal language of $\mathcal{E}$, we can build the *internal subset relation*

$$[\![C : \mathcal{P}_{\mathrm{s}}(X).C' : \mathcal{P}_{\mathrm{s}}(X).\forall x : X(x \in_X C) \to (x \in_X C')]\!] \to \mathcal{P}_{\mathrm{s}}(X) \times \mathcal{P}_{\mathrm{s}}(X).$$

We will often denote this relation by $\subseteq_X \to \mathcal{P}_{\mathrm{s}}(X) \times \mathcal{P}_{\mathrm{s}}(X)$. Requiring the internal subset relation to be small, is a last axiom imposed on the small maps. We will later see that this axiom corresponds to the powerset axiom of set theory.

(P2) For every object X, the internal subset relation is small.

We finish this section by providing the definition of a category of classes.

Definition 2.2.3. We call a Heyting pretopos $\mathcal{E}$ equipped with a class $\mathcal{S}$ satisfying axioms (S1)-(S5), (P1) and (P2) and that has a stable and small natural number object is called a *category of classes.*

Note that in the definition of a category of classes we also required the natural numbers object to be small. As we will see later, this requirement corresponds to the axiom of infinity. It should also be remarked that the our axiomatization of the small maps is in fact redundant. In [28] for example Simpson shows that axiom (S3) is implied by the other axioms we required. We will make use of Simpson's observation in the next chapter when we examine examples of categories of classes.

2.3 Properties of Categories of Classes

In this section we will derive some elementary properties of a categories of classes that we will use later in this thesis. First, we will derive more consequences of the axioms for small maps. Then we will examine the representability axiom (P1) more closely. In particular, we will see that the assignment $X \mapsto \mathcal{P}_{\mathrm{s}}(X)$ is functorial giving rise to the generalized powerclass functor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$. We will then discuss that the notion of a category of classes is stable under slicing. This property allows us to reason in the internal logic also when formulas contain free parameters. We will see that an initial algebra for the $\mathcal{P}_{\mathrm{s}}$-functor is also stable under slicing in a certain sense. Then we will discuss another interesting property of a category of classes, namely that the small objects in category of classes form a topos. This property shows us that the small objects in a Heyting pretopos indeed behave like sets. Finally, we will show that small objects in a category of classes are exponentiable.

2.3.1 Properties of the Small Maps

Next, we will examine some consequences of the axioms for small maps. The first two properties in the following Lemma state that small maps are preserved by the product and coproduct functor, respectively. The last property is called cancellation in [28]. From this property it is immediate to see that all arrows between small objects are small.

Lemma 2.3.1. *Assume $f : A \to X$ and $g : B \to Y$ are arrows in $\mathcal{E}$.*

(1) *If f and g are small, the product $f \times g : A \times B \longrightarrow X \times Y$ is a small map.*

(2) *If f and g are small also $f + g : A + B \longrightarrow X + Y$ is small.*

(3) *Let $f : A \to B$ and $g : B \to C$ be arrows in $\mathcal{E}$. If $h := g \circ f$ is small then also f is small.*

Proof. **(1):** By (S4) it follows that f' and g' in the pullbacks below are small.

$$\begin{tikzcd} A \times Y \arrow[r] \arrow[d, "f'"'] & A \arrow[d, "f"] \\ X \times Y \arrow[r, "\pi_X"'] & X \end{tikzcd} \qquad \begin{tikzcd} B \times X \arrow[r] \arrow[d, "g'"'] & B \arrow[d, "g"] \\ X \times Y \arrow[r, "\pi_Y"'] & Y \end{tikzcd}$$

Again, by (S4) it follows that in the pullback below p is small.

$$\begin{tikzcd} B \times A \arrow[r] \arrow[d, "p"'] & A \times Y \arrow[d, "f'"] \\ B \times X \arrow[r, "g'"'] & X \times Y \end{tikzcd}$$

It is easy to see that $f \times g = g' \circ p$ so (1) follows since small maps are closed under composition.

(2): Since in the Heyting pretopos coproducts are disjoint, the coproduct inclusion $m_A : A \to A + B$, $m_B :\to A + B$ are monos, so in particular small by axiom (S1). So by (S2), the maps $m_A \circ f : X \to A + B$ and $m_B \circ g : Y \to A + B$ are small. Now the claim follows by axiom (S5).

(3): The map h' in the pullback

$$\begin{tikzcd} D \arrow[r] \arrow[d, "h'"'] & A \arrow[d, "h"] \\ B \arrow[r, "g"'] & C \end{tikzcd}$$

is small. Since $g \circ f = h$ there is a unique map $l : A \to D$ such that $f = h' \circ l$. It is clear that l is mono and therefore small, so again the claim follows by (S2).

□

Lemma 2.3.2. (1) *Assume $s : S \to X \times Y$ is a relation and $r : R \to X \times Y$ is a relation that is small in X. Then the intersection $R \cap S \to X \times Y$ is also small in X.*

(2) *Similarly, assume $R \to Y \times Z$ is a small in Y and $S \to X \times Y$ is a small in X. Let R' and S' be the pullbacks in*

$$\begin{tikzcd} R' \arrow[r] \arrow[d] & R \arrow[d] \\ X \times Y \times Z \arrow[r] & Y \times Z \end{tikzcd} \qquad \begin{tikzcd} S' \arrow[r] \arrow[d] & S \arrow[d] \\ X \times Y \times Z \arrow[r] & X \times Y \end{tikzcd}$$

Then the intersection $R' \cap S' :\to (X \times Y) \times Z$ is small in $X \times Y$

Proof. The intersection of R and S is given by the pullback

$$\begin{tikzcd} R \cap S \arrow[r, "s'"] \arrow[d] & R \arrow[d, "r"] \\ S \arrow[r, "s"'] & X \times Y \end{tikzcd}.$$

Since s is a mono and therefore small, it follows that s' is small. Now the composite

$$R \cap S \xrightarrow{r \circ s'} X \times Y \xrightarrow{\pi_Y} Y$$

is small since $\pi_Y \circ r$ is small since R is small in X. This proves (1). The proof of (2) is similar. □

Next, we will examine axiom (P1) a bit more closely following a discussion in [26]. The axiom states that there is a one-to-one correspondence between small relations on $X \times Y$ and arrows from Y to $\mathcal{P}_s(X)$. Given a small relation $r : R \to X \times Y$, write $\lfloor r \rfloor : Y \to \mathcal{P}_s(X)$ for the unique map classifying r. Conversely, for every function $f : Y \to \mathcal{P}_s(X)$ write $\lceil f \rceil := (\mathrm{id}_X \times f)^{-1}(\in_X)$ in $\mathrm{Sub}(X \times Y)$ for the unique small relation classified by f. Observe that in the internal logic we have for every small relation $R \to X \times Y$

$$\mathcal{E} \models_{x:X,y:Y} Rxy \leftrightarrow x \in_X \lfloor r \rfloor(y).$$

And conversely, for every arrow $f : Y \to \mathcal{P}_s(X)$

$$\mathcal{E} \models_{x:X,y:Y} x \in_X f(y) \leftrightarrow \lceil f \rceil xy.$$

Next we will show that the element relation $\in_X$ is extensional. This property will turn out to be very useful when arguing internally about $\mathcal{E}$.

Lemma 2.3.3. Typed Extensionality:

$$\mathcal{E} \models_{A:\mathcal{P}_s(X),B:\mathcal{P}_s(X)} (\forall x : X.(x \in_X A \leftrightarrow x \in_X B)) \leftrightarrow A = B.$$

Proof. Let ϕ abbreviate the formula $(\forall x : X.(x \in_X A \leftrightarrow x \in_X B))$. Let

$$T := [\![A : \mathcal{P}_s(X)B : \mathcal{P}_s(X).\phi]\!] \xrightarrow{\langle r_1, r_2 \rangle} \mathcal{P}_s(X) \times \mathcal{P}_s(X)$$

be the interpretation of ϕ. In order to show that

$$\mathcal{E} \models_{A:\mathcal{P}_s(X),B:\mathcal{P}_s(X)} (\forall x : X.(x \in_X A \leftrightarrow x \in_X B)) \to A = B. \qquad (2.2)$$

we need to show that $r = \langle r_1, r_2 \rangle$ factors through the diagonal $\Delta_{\mathcal{P}_s(X)} : \mathcal{P}_s(X) \to \mathcal{P}_s(X) \times \mathcal{P}_s(X)$. To this end let U be the pullback in

$$\begin{tikzcd}
U \arrow[r] \arrow[d] & \in_X \times \mathcal{P}_{\mathrm{s}}(X) \arrow[r] \arrow[d] & \in_X \arrow[d] \\
X \times T \arrow[r, "\mathrm{id}_X \times r"] & X \times \mathcal{P}_{\mathrm{s}}(X) \times \mathcal{P}_{\mathrm{s}}(X) \arrow[r, "\langle \pi_1, \pi_2 \rangle"] & X \times \mathcal{P}_{\mathrm{s}}(X)
\end{tikzcd}$$

so that r_1 classifies U. Similarly let V be the pullback in

$$\begin{tikzcd}
V \arrow[r] \arrow[d] & \in_X \times \mathcal{P}_{\mathrm{s}}(X) \arrow[r] \arrow[d] & \in_X \arrow[d] \\
X \times T \arrow[r, "\mathrm{id}_X \times r"] & X \times \mathcal{P}_{\mathrm{s}}(X) \times \mathcal{P}_{\mathrm{s}}(X) \arrow[r, "\langle \pi_1, \pi_3 \rangle"] & X \times \mathcal{P}_{\mathrm{s}}(X)
\end{tikzcd}$$

so that V is classified by r_2. Now using the internal logic we can prove that

$$U = [\![\phi \wedge x \in_X A]\!] = [\![\phi \wedge x \in_X B]\!] = V \text{ in } \mathrm{Sub}(X \times \mathcal{P}_{\mathrm{s}}(X) \times \mathcal{P}_{\mathrm{s}}(X)).$$

So U and V define in fact the same subobject of $X \times T$. By the uniqueness of the classifying map it follows that $r_1 = r_2$. Therefore, $\langle r_1, r_2 \rangle$ factors through the diagonal as desired and we conclude that (2.2) holds. The proof of the other direction is even simpler. $\square$

The next lemma states that the assignment $X \mapsto \mathcal{P}_{\mathrm{s}}(X)$ given by $(P1)$ is functorial. An arrow $f : A \to B$ will be mapped to $\mathcal{P}_{\mathrm{s}}(f) : \mathcal{P}_{\mathrm{s}}(A) \to \mathcal{P}_{\mathrm{s}}(B)$, where $\mathcal{P}_{\mathrm{s}}(f)$ maps a subset of A to its image under f. For a category of classes $\mathcal{E}$, we will usually refer to $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$ as the (generalized) powerclass functor.

Lemma 2.3.4. *The assignment $A \mapsto \mathcal{P}_{\mathrm{s}}(A)$ extends to a functor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$.*

Proof. Let $f : Y \to X$ be an arrow $\mathcal{E}$. The composite

$$\in_Y \xrightarrow{m_Y} Y \times \mathcal{P}_{\mathrm{s}}(Y) \xrightarrow{\langle f\pi_Y, \pi_{\mathcal{P}_{\mathrm{s}}(Y)}, \pi_Y \rangle} X \times \mathcal{P}_{\mathrm{s}}(Y) \times Y \; .$$

is easily seen to define the subobject $[\![y \in_Y A \wedge f(y) = x]\!]$ in $\mathrm{Sub}(X \times \mathcal{P}_{\mathrm{s}}(Y) \times Y)$, where the free variables x, y and A are of type X, Y and $\mathcal{P}_{\mathrm{s}}(Y)$, respectively.

Therefore, the image m in the diagram

$$\begin{tikzcd}
\in_Y \arrow[r, tail, "m_Y"] \arrow[dr, two heads, "e"'] & Y \times \mathcal{P}_{\mathrm{s}}(Y) \arrow[r, "f \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(Y)}"] & X \times \mathcal{P}_{\mathrm{s}}(Y) \\
& \mathrm{Im}\big(f \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m_Y\big) \arrow[ur, tail, "m"'] &
\end{tikzcd}$$

defines the subobject $[\![\exists y : Y(y \in_Y A \wedge f(y) = x))]\!]$ in $\mathrm{Sub}(X \times \mathcal{P}_{\mathrm{s}}(Y))$. We will now argue that m defines a small relation on $X \times \mathcal{P}_{\mathrm{s}}(Y)$.

We have $\pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m \circ e = \pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m_Y \circ f \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(Y)} = \pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m_Y$. By (P1) it follows that $\pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m_Y$ is small and therefore $\pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m \circ e$ is small. Now since e is epi, by (S3) it follows that $\pi_{\mathcal{P}_{\mathrm{s}}(Y)} \circ m$ is small. So m indeed defines a small relation on $X \times \mathcal{P}_{\mathrm{s}}(Y)$.

By (P1) we define $\mathcal{P}_{\mathrm{s}}(f)$ as the classifying map of the relation m, i.e. the unique map fitting in the pullback

$$\begin{tikzcd} \mathrm{Im}\big(f \times \mathrm{id}_{\mathcal{P}_\mathrm{s}(Y)} \circ m_Y\big) \arrow[r] \arrow[d, tail] & \in_X \arrow[d, tail] \\ X \times \mathcal{P}_\mathrm{s}(Y) \arrow[r, "\mathrm{id}_X \times \mathcal{P}_\mathrm{s}(f)"] & X \times \mathcal{P}_\mathrm{s}(X) \end{tikzcd} .$$

Note, that by the above we have that

$$\mathcal{E} \models_{x:X, A:\mathcal{P}_\mathrm{s}(Y)} x \in_X \mathcal{P}_\mathrm{s}(f)(A) \leftrightarrow \exists y : Y(y \in_Y A \wedge f(y) = x). \tag{2.3}$$

Using (2.3) and typed extensionality it is easy to verify that the assignment $f \mapsto \mathcal{P}_\mathrm{s}(f)$ is functorial. □

Observe that (2.3) in the proof of Lemma 2.3.4 expresses in the internal logic that for $f : Y \to X$, $\mathcal{P}_\mathrm{s}(f)$ maps a subset A of Y to the image of A under f.

Corollary 2.3.5. *The functor $\mathcal{P}_\mathrm{s} : \mathcal{E} \to \mathcal{E}$ preserves monos.*

Proof. This is easy to see using the internal logic and the characterization (2.3) in the proof of Lemma 2.3.4. □

Lemma 2.3.6. *If $f : A \to X$ is small, we have an internal inverse image map $f^{-1} : X \to \mathcal{P}_\mathrm{s}(A)$ such that*

$$\mathcal{E} \models_{x:X, a:A} a \in_A f^{-1}(x) \leftrightarrow f(a) = x$$

Proof. Since f is small, the graph $\langle \mathrm{id}_A, f\rangle : A \to A \times X$ of f defines a small relation. The classifying map of this relation, i.e. the unique map $f^{-1} : X \to \mathcal{P}_\mathrm{s}(A)$ fitting in the pullback

$$\begin{tikzcd} A \arrow[r] \arrow[d, "\langle \mathrm{id}_A, f\rangle"'] & \in_A \arrow[d] \\ A \times X \arrow[r, "\mathrm{id}_A \times f^{-1}"'] & A \times \mathcal{P}_\mathrm{s}(A) \end{tikzcd}$$

satisfies the above condition. □

As a well known fact in topos theory, there is also a contravariant version of the powerset functor. This would act on $f : A \to B$ by mapping a subset of A to its preimage under f. In our context this applies only to small maps.

Lemma 2.3.7. *If $f : A \to B$ is small, there is a map $f' : \mathcal{P}_\mathrm{s}(B) \to \mathcal{P}_\mathrm{s}(A)$ such that*

$$\mathcal{E} \models z \in_A f'(C) \leftrightarrow \exists y : By \in_B C \wedge f(z) = y.$$

Proof. Let $f : A \to B$ be a small map in $\mathcal{E}$. We define $f' : \mathcal{P}_\mathrm{s}(B) \to \mathcal{P}_\mathrm{s}(A)$ as follows. Take R to be the relation $R := [\![z : A, x : \mathcal{P}_\mathrm{s}(B) \mid \exists y : B(y \in_B x \wedge f(z) = y)]\!]$. We show that this defines a small relation on $A \times \mathcal{P}_\mathrm{s}(B)$. Consider the following diagram.

$$
\begin{tikzcd}
T \arrow[r] \arrow[d, "h"] \arrow[dd, bend right, "e"'] & A\times \in_B \arrow[r] \arrow[d] & \in_B \arrow[d] \\
A \times \mathcal{P}_{\mathrm{s}}(B) \arrow[r, "\langle \pi_1, f\pi_1, \pi_2\rangle"] & A \times B \times \mathcal{P}_{\mathrm{s}}(B) \arrow[r] \arrow[d, "\langle \pi_1, \pi_3\rangle"] & B \times \mathcal{P}_{\mathrm{s}}(B) \arrow[d, "\pi_2"] \\
R \arrow[r, tail, "m"] & A \times \mathcal{P}_{\mathrm{s}}(B) \arrow[r, "p"] & \mathcal{P}_{\mathrm{s}}(B)
\end{tikzcd},
$$

where the upper right rectangle is a pullback and me is the epi mono factorization of $\langle \pi_1, \pi_3\rangle\langle \pi_1, f\pi_1, \pi_2\rangle \circ h$. To show that R is small in A, we need to show that pm is a small map. The horizontal map from $A \times \mathcal{P}_{\mathrm{s}}(B)$ to $B \times \mathcal{P}_{\mathrm{s}}(B)$ is $f \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(B)}$ and therefore small. It follows by (S4) that also the composite on the top of the rectangle is small. The map map on the right hand side of the rectangle is small since $\in_B$ is a small relation. As a composite of small maps pme is small. Now since e is epi, by (S3) it follows that pm is small as desired. We define f' as the classifying map of this relation. Note that

$$\mathcal{E} \models z \in_A f'(C) \leftrightarrow \exists y : By \in_B C \wedge f(z) = y.$$

as required. □

It is interesting to observe that the additional structure of small maps imposes logical properties on the Heyting pretopos. In particular a category of classes always has a subobject classifier.

Lemma 2.3.8. *$\mathcal{E}$ has a subobject classifier.*

Proof. We claim that $\in_1 \to \mathcal{P}_{\mathrm{s}}(1)$ is a subobject classifier. Let $m : A \to X$ be a mono in $\mathcal{E}$. Since monos are small, m gives rise to a small relation $m : A \to X \cong 1 \times X$. By (P1) m corresponds to a unique map $\lfloor m \rfloor$ fitting in the pullback

$$
\begin{tikzcd}
A \arrow[r] \arrow[d, tail, "m"'] & \in_1 \arrow[d, tail] \\
X \arrow[r, "\lfloor m \rfloor"'] & \mathcal{P}_{\mathrm{s}}(1)
\end{tikzcd}
$$

Note that one can infer that $\in_1$ is isomorphic to the terminal object of $\mathcal{E}$. □

A useful property that we will use later and that we stole from [30] is described in the following lemma.

Lemma 2.3.9. *If $m : A \to B$ is a mono, then the following square is a pullback:*

$$
\begin{tikzcd}
\in_A \arrow[r] \arrow[d] & \in_B \arrow[d] \\
A \times \mathcal{P}_{\mathrm{s}}(A) \arrow[r, "m \times \mathcal{P}_{\mathrm{s}}(m)"'] & B \times \mathcal{P}_{\mathrm{s}}(B)
\end{tikzcd}
$$

Proof. Since m is a mono, the composite

$$\in_A \rightarrowtail A \times \mathcal{P}_{\mathrm{s}}(A) \xrightarrow{m \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(A)}} B \times \mathcal{P}_{\mathrm{s}}(A)$$

provides its own epi-mono factorization. According to the proof of Lemma 2.3.4, map $\mathcal{P}_{\mathrm{s}}(m)$ is then defined to be the classifying map of this relation, i.e. the one fitting in the pullback

$$\begin{tikzcd}
\in_A \arrow[rr] \arrow[d] & & \in_B \arrow[dd] \\
A \times \mathcal{P}_{\mathrm{s}}(A) \arrow[d, "m \times \mathrm{id}_{\mathcal{P}_{\mathrm{s}}(A)}"'] & & \\
B \times \mathcal{P}_{\mathrm{s}}(A) \arrow[rr, "\mathrm{id}_B \times \mathcal{P}_{\mathrm{s}}(m)"'] & & B \times \mathcal{P}_{\mathrm{s}}(B)
\end{tikzcd}$$

Now the claim follows immediately. □

2.3.2 Stability under Slicing

In this section, we will discuss that the notion of a category of classes is stable under the formation of slice categories.

Proposition 2.3.10. *If $\mathcal{E}$ is a category of classes, then for every object X the slice category $\mathcal{E}/X$ is a category of classes. Moreover, for every $f : Y \to X$, the pullback functor $f^* : \mathcal{E}/X \to \mathcal{E}/Y$ preserves all the structure in question.*

A detailed proof of the above proposition can for example be found in [30]. We will sketch the basic concepts. Given an object X of $\mathcal{E}$ it is not difficult to see that the slice category $\mathcal{E}/X$ is a Heyting pretopos. We equip $\mathcal{E}/X$ with a class of small maps given by

$$\mathcal{S}^X := \{s : (g : A \to X) \to (h : B \to X) \mid s : A \to B \text{ is small in } \mathcal{E}\},$$

so we define a map in the slice category to be small if and only if it is a small map in $\mathcal{E}$. It is easy to see that $\mathcal{S}^X$ satisfies axioms $(S1) - (S5)$. Given an object $g : A \to X$ in $\mathcal{E}/X$ let

$$T := [\![C : \mathcal{P}_{\mathrm{s}}(A), x : X. \forall a : A(x \in_A C \to g(a) = x)]\!] \rightarrowtail^{t} \mathcal{P}_{\mathrm{s}}(A) \times X \;.$$

Then the powerclass $\mathcal{P}_s^X(g : A \to X)$ of $g : A \to X$ can be defined as the composite

$$T \overset{t}{\rightarrowtail} \mathcal{P}_{\mathrm{s}}(A) \times X \xrightarrow{\pi_X} X \;.$$

In order to define the element relation in the slice category, let

$$S := [\![\forall a : A(x \in_A C \to g(a) = x) \wedge a \in_A C]\!] \overset{s}{\rightarrowtail} A \times \mathcal{P}_{\mathrm{s}}(A) \times X \;,$$

where of course the free variables are of types $a : A, C : \mathcal{P}_{\mathrm{s}}(A)$ and $x : X$, respectively. Then $\in_{g:A\to X}$ is defined as the composite

$$S \overset{s}{\rightarrowtail} A \times \mathcal{P}_{\mathrm{s}}(A) \times X \xrightarrow{\pi_X} X \;.$$

A detailed proof of the validity of axiom (P1) is given in [30]. That also the powerset axiom (P2) is valid in the slice category is for example proven in [26]. Concerning the preservation under pullback, it is easy to see that the Heyting pretopos structure is preserved and it follows by axiom (S4), that also the class of small maps is preserved by the pullback functors. For a proof that powerclasses and the element relation are preserved by the pullback functors, the reader is referred to the proof of Lemma 2.1.16 of [30].

2.3.3 The Initial Algebra for the $\mathcal{P}_{\mathrm{s}}$ Functor

We will now discuss that also the structure of an initial algebra for the $\mathcal{P}_{\mathrm{s}}$ functor is stable, i.e. if $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$ is the initial algebra for $\mathcal{P}_{\mathrm{s}}$ in $\mathcal{E}$ then for every T of $\mathcal{E}$, $t^*(\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V)$ is the initial algebra for ${\mathcal{P}_s}^T$ in the slice category $\mathcal{E}/T$, where t^* is the pullback functor associated to the unique map $t : T \to 1$.

We say that a $\mathcal{P}_{\mathrm{s}}$ algebra has the *inductive property* if it has no proper subalgebras. It is easy to see that an initial algebra for the $\mathcal{P}_{\mathrm{s}}$-functor satisfies the inductive property. For if $\eta : \mathcal{P}_{\mathrm{s}}(V) \to V$ is initial and $\alpha : \mathcal{P}_{\mathrm{s}}(A) \to A$ is a subalgebra then by initiality there is a unique algebra homomorphism h fitting in the diagram

$$\begin{tikzcd} \mathcal{P}_{\mathrm{s}}(A) \arrow[r, "\mathcal{P}_{\mathrm{s}}(g)"] \arrow[d, "\alpha"] & \mathcal{P}_{\mathrm{s}}(V) \arrow[r, "\mathcal{P}_{\mathrm{s}}(h)"] \arrow[d, "\eta"] & \mathcal{P}_{\mathrm{s}}(A) \arrow[d, "\alpha"] \\ A \arrow[r, "g"] & V \arrow[r, "h"] & A \end{tikzcd}$$

Since the identity is the only algebra homomorphism on $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$, we conclude that $gh = \mathrm{id}_V$ and since g is a mono it follows that $hg = \mathrm{id}_A$, showing that g is an isomorphism of $\mathcal{P}_{\mathrm{s}}$-algebras. One particular property of the $\mathcal{P}_{\mathrm{s}}$-functor is that a $\mathcal{P}_{\mathrm{s}}$-algebra that is also a fixed point of $\mathcal{P}_{\mathrm{s}}$ satisfying the inductive property is automatically initial. This can be proved by imitating the proof of the recursion theorem familiar from set theory in the categorical setting.

Proposition 2.3.11. *Assume $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$ is a fixed point of $\mathcal{P}_{\mathrm{s}}$ having the inductive property. Then $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$ is an initial algebra.*

A similar result is discussed in [5, p. 19]. For the remainder of this section, we will summarize how the initial algebra of the $\mathcal{P}_{\mathrm{s}}$ functor behaves under slicing. So let T be an object on $\mathcal{E}$ and as above $\eta : \mathcal{P}_{\mathrm{s}}(V) \to V$ the initial algebra for $\mathcal{P}_{\mathrm{s}}$. Note that since $\mathcal{P}_{\mathrm{s}}$ is indexed by Proposition 2.3.10, we can pull $\eta : \mathcal{P}_{\mathrm{s}}(V) \to V$ back along $t : T \to 1$ and obtain a ${\mathcal{P}_s}^T$-algebra $t^*(\mathcal{P}_{\mathrm{s}}(V) \to V)$ in $\mathcal{E}/T$. It is immediate to see that $t^*(\mathcal{P}_{\mathrm{s}}(V) \to V)$ is a fixed point of ${\mathcal{P}_{\mathrm{s}}}^T$. One can also prove that it still has the inductive property. So we obtain:

Lemma 2.3.12. *Assume $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$ is a fixed point of V having the inductive property. Then $t^*(\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V)$ has the inductive property.*

Note that together with the characterization stated in Proposition 2.3.11, we obtain that initial algebras for the $\mathcal{P}_{\mathrm{s}}$ functor are indexed.

Corollary 2.3.13. *Assume, $\alpha : \mathcal{P}_{\mathrm{s}}(V) \to V$ is the initial algebra for $\mathcal{P}_{\mathrm{s}}$ in $\mathcal{E}$. Then for every T in $\mathcal{E}$, $\langle \alpha, \mathrm{id}_T \rangle : {\mathcal{P}_{\mathrm{s}}}^T(t^*(V)) \cong \mathcal{P}_{\mathrm{s}}(V) \times T \to V \times T$ is the initial algebra for ${\mathcal{P}_s}^T : \mathcal{E}/T \to \mathcal{E}/T$.*

2.3.4 The Topos of Small Subobjects

We will now see that the full subcategory of small objects in a category of classes carries the structure of a topos. This was pointed out for example in [2, 26, 30]. This category will be denoted by $\mathcal{E}_{\mathcal{S}}$, it is given by:

- The objects of $\mathcal{E}_{\mathcal{S}}$ are the small objects of $\mathcal{E}$.
- The arrows of $\mathcal{E}_{\mathcal{S}}$ are the arrows $f : A \to B$ in $\mathcal{E}$ between small objects of $\mathcal{E}$.

In other words, $\mathcal{E}_{\mathcal{S}}$ is the full subcategory of $\mathcal{E}$ consisting of the small objects. It follows by (3) in Lemma 2.3.1 that all arrows of $\mathcal{E}_{\mathcal{S}}$ are in fact small in $\mathcal{E}$.

Lemma 2.3.14. *Let $\mathcal{E}$ be a category of classes. Then for every small object A of $\mathcal{E}$ also $\mathcal{P}_{\mathrm{s}}(A)$ is small.*

Proof. We show more generally, that for every small object $g : A \to X$ of $\mathcal{E}\,/X$ also $\mathcal{P}_s^X(g : A \to X)$ is a small object of $\mathcal{E}\,/X$. So assume that $g : A \to X$ is a small object of $\mathcal{E}\,/X$. But this just means that the arrow $g : A \to X$ is a small map in $\mathcal{E}$. Now using Lemma 2.3.6, there is a map $g^{-1} : X \to \mathcal{P}_{\mathrm{s}}(A)$ such that

$$\mathcal{E} \models \forall a : A(a \in_A g^{-1}(x) \leftrightarrow g(a) = x).$$

Using the internal logic it easy to see that T, the domain of $\mathcal{P}_s^X(g : A \to X)$ can alternatively be obtained by the above square in the pullback

$$\begin{tikzcd}
T \arrow[rr] \arrow[d] & & \subseteq \arrow[d] \\
\mathcal{P}_{\mathrm{s}}(A) \times X \arrow[rr, "\mathrm{id}_{\mathcal{P}_{\mathrm{s}}(A)} \times g^{-1}"] \arrow[d] & & \mathcal{P}_{\mathrm{s}}(A) \times \mathcal{P}_{\mathrm{s}}(A) \arrow[d] \\
X \arrow[rr, "g^{-1}"] & & \mathcal{P}_{\mathrm{s}}(A)
\end{tikzcd} .$$

By axiom (P2) the map on the right hand side is small, so this shows by (S4) that $\mathcal{P}_s^X(g : A \to X)$ is a small object in $\mathcal{E}\,/X$.

$\square$

Now it is easy to show that $\mathcal{E}_{\mathcal{S}}$ is an elementary topos.

Proposition 2.3.15. *The category $\mathcal{E}_{\mathcal{S}}$ is an elementary topos and all the relevant structure is inherited from $\mathcal{E}$.*

Proof. Similarly as in Lemma 2.3.1 one can see that products of small objects are small. Moreover, if $f, g : A \to B$ are two maps in $\mathcal{E}_{\mathcal{S}}$, then their equalizer in $\mathcal{E}$ is small and therefore provides the equalizer of f and g in $\mathcal{E}_{\mathcal{S}}$. By Lemma 2.3.14, the subobject classifier $\mathcal{P}_{\mathrm{s}}(1)$ is small and provides a subobject classifier in $\mathcal{E}_{\mathcal{S}}$. Again, using Lemma 2.3.14, for every small object A also $\mathcal{P}_{\mathrm{s}}(A)$ is small and $\in_A$ is a small relation on $A \times \mathcal{P}_{\mathrm{s}}(A)$ in $\mathcal{E}_{\mathcal{S}}$. Since every relation on $A \times B$ for some B in $\mathcal{E}_{\mathcal{S}}$ is small, axiom (P1) expresses precisely that $\in_A \rightarrow A \times \mathcal{P}_{\mathrm{s}}(A)$ satisfies the universal property of powerobjects. $\square$

2.3.5 Exponentials and the Object of Epimorphisms

In this last section, we will see that small objects in a category of classes are exponentiable. Moreover, we will introduce the object of epimorphisms. Similar, as the exponential serves as a categorical notion of a function space, the object of epimorphisms represents the collection of surjective functions between two objects. The following proofs are taken from [30].

Let A be a small object of $\mathcal{E}$. For an arbitrary object B we define

$$B^A := \{f : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists! y : B (x, y) \in_{A\times B} f\}.$$

Lemma 2.3.16. *The following correspondence holds in* $\mathcal{E}$

$$\frac{A \to B}{\overline{\overline{}}\; 1 \to B^A},$$

i.e. arrows from A *to* B *correspond to points of* B^A*.*

Proof. Let $g : 1 \to B^A$ be given. Let $\lfloor g \rfloor := g \circ m$, where $m : B^A \to \mathcal{P}_{\mathrm{s}}(A \times B)$. Since $\lfloor g \rfloor$ factors through m, it follows that $[\![\forall x : A \exists! b : B (x, y) \in_{A\times B} \lfloor g \rfloor]\!] \cong 1$. Therefore, the relation $[\![(x, y) \in_{A\times B} \in \lfloor g \rfloor]\!] \to A \times B$ is functional. By Lemma 2.1.16 we conclude that it corresponds to a unique map $g' : A \to B$ such that

$$\Gamma(g') = [\![g'(x) = y]\!] = [\![(x, y) \in_{A\times B} \lfloor g \rfloor]\!] \text{ in } \mathrm{Sub}(A \times B).$$

Conversely, assume that $f : A \to B$ is an arrow in $\mathcal{E}$. Let $\langle \mathrm{id}_A, f\rangle \cong \Gamma(f) \to A \times B$ be the graph of f. Since A is small, the composite

$$A \xrightarrow{\langle \mathrm{id}_A, f\rangle} A \times B \longrightarrow 1$$

is a small map. But this shows that the graph of f is a small subobject of $A \times B$. Let $\lceil f \rceil : 1 \to \mathcal{P}_{\mathrm{s}}(A \times B)$ be its classifying map according to (P1). Since $[\![(x, y) \in_{A\times B} \lceil f \rceil]\!] \cong [\![f(x) = y]\!]$ it is easy to se that $[\![\forall x : A \exists! y : B (x, y) \in_{A\times B} \lceil f \rceil]\!] \cong 1$. Therefore, $\lceil f \rceil$ factors through B^A via an arrow $f' : 1 \to B^A$. One can easily check that the concepts are inverse to each other. □

Proposition 2.3.17. *For every object* E *of* $\mathcal{E}$ *we have*

$$\frac{E \times A \to B}{E \to B^A}.$$

Proof. Let $e : E \to 1$ be the unique map from E to the terminal object and $e^* : \mathcal{E} \to \mathcal{E}/E$ the corresponding pullback functor. Then $e^*(A \to 1) = \pi_1 : E \times A \to E$ and $e^*(B \to 1) = \pi_1 : E \times B \to E$ in $\mathcal{E}/E$. Let us denote these objects by e_A and e_B, respectively. Since we showed that pullback functors preserve the relevant structure, we have that $E \times B^A \to E = e^*(B^A \to 1)$ represents the subobject $[\![z : {\mathcal{P}_{\mathrm{s}}}^E(e_A \times_E e_B) \mid \forall x : e_A \exists! y : e_B \, B(x, y) \in^E_{e_A \times_E e_B} z]\!]$ of ${\mathcal{P}_{\mathrm{s}}}^E(e_A \times_E e_B)$. So we can use the result from Lemma 2.3.16 for e_A and e_B in the slice category $\mathcal{E}/E$, to obtain a one to one correspondence between the maps on the left and the right hand side.

$$
\begin{tikzcd}
E \times A \arrow[rr, "l"] \arrow[dr] & & E \times B \arrow[dl] \\
& E &
\end{tikzcd}
\qquad
\begin{tikzcd}
E \arrow[rr, "h"] \arrow[dr] & & E \times B^A \arrow[dl] \\
& E &
\end{tikzcd}
$$

It is easy to see that the map l and h above need to be of the form $l = \langle \pi_E, f\rangle$ for some $f : E\times A \to B$ and $h = \langle \mathrm{id}_E, g\rangle$ for some $g : E \to B^A$ in $\mathcal{E}$, respectively. This establishes the desired correspondence. $\square$

This finishes the proof that small objects in $\mathcal{E}$ are exponentiable. Now similarly, if A is a small and B an arbitrary object of $\mathcal{E}$, we define

$$\mathbf{Epi}(A, B) := \{f : \mathcal{P}_{\mathrm{s}}(A\times B) \mid \forall x : A\exists! y : B\langle x, y\rangle \in_{A\times B} f \wedge \forall y : B\exists x : A\langle x, y\rangle \in_{A\times B} f\},$$

the *object of epimorphisms from A to B.*

Proposition 2.3.18. *There is a one to one correspondence*

$$\frac{E \times A \to B \textit{ epi in } \mathcal{E}}{\overline{E \to \mathbf{Epi}(A, B)}} \quad .$$

Proof. Again, we first show that there is a one to one correspondence

$$\frac{A \to B \text{ epi in } \mathcal{E}}{\overline{1 \to \mathbf{Epi}(A, B)}} \quad \text{in } \mathcal{E}.$$

Let $g : 1 \to \mathbf{Epi}(A, B)$ is an arrow in $\mathcal{E}$ and let $\lfloor g\rfloor$ be the composite

$$1 \xrightarrow{g} \mathbf{Epi}(A, B) \rightarrowtail B^A \rightarrowtail \mathcal{P}_{\mathrm{s}}(A \times B) \; .$$

As $\lfloor g\rfloor$ factors through B^A there is a unique map $g' : A \to B$ such that $[\![g'(x) = y]\!] \cong [\![(x, y) \in_{A\times B} \lfloor g\rfloor]\!]$. Since $\lfloor g\rfloor$ factors trough $\mathbf{Epi}(A, B)$ one can easily infer that $[\![\forall y : B\exists x : Ag'(x) = y]\!] \cong 1$. In other words g' is a cover and therefore epi.

Conversely, assume $f : A \to B$ is epi. and let $\lceil f\rceil : 1 \to \mathcal{P}_{\mathrm{s}}(A \times B)$ be the classifying map of the graph of f as above. Since f is epi, f is a cover by (5) in Proposition 2.1.3. Using this it is easy to see that $\lceil f\rceil$ factors through $\mathbf{Epi}(A, B)$. We obtain the correspondence claimed in the proposition by passing to the suitable slice category as above. All arguments are as above, one only needs to observe that an arrow $g : E \times A \to B$ is epi in $\mathcal{E}$ if and only if $h := \langle \mathrm{id}_E, g\rangle : e_A \to e_B$ is epi in $\mathcal{E}/E$. $\square$

Proposition 2.3.19 clearly justifies the terminology of the object of epimorphisms. A property that we will use later can also be found in [20, p. 287].

Lemma 2.3.19. *Let $\mathcal{E}$ be a Boolean pretopos , $m : Z \to Y$ be a mono and $z : 1 \to Z$ be a global section. Then $\mathbf{Epi}(X, Z) \cong 0$ implies $\mathbf{Epi}(X, Y) \cong 0$.*

Proof. By the correspondence in Proposition 2.3.18 the identity map on $\mathbf{Epi}(X, Y)$ gives rise to an epi $e : \mathbf{Epi}(X, Y) \times X \to Y$. Since $\mathcal{E}$ is Boolean $m : Z \to Y$ has a Boolean complement $m' : Z' \to Y$ in $\mathrm{Sub}(Y)$, i.e. $Y = Z + Z'$. The maps id_Z and $z\circ !_{Z'}$ give rise to an arrow $r : Y \to Z$ with $\mathrm{id}_Z = rm$. By the latter, r is epi. So the composite $r \circ e : \mathbf{Epi}(X, Y) \times X \to Z$ is epi. So by using the correspondence in Proposition 2.3.18, $r \circ e$ corresponds to an arrow $\mathbf{Epi}(X, Y) \to \mathbf{Epi}(X, Z) \cong 0$. Since in a boolean pretopos every arrow to 0 is an isomorphism, the claim follows. $\square$

Chapter 3

Examples of Categories of Classes

Having developed the general theory of categories of classes in the previous chapter, we will now turn to examine examples of such. First, we will indicate how the familiar classes of set theory fit in our framework. Then we will discuss the categories of large presheaves and the category of large sheaves on site, the class-sized version of presheaves and sheaves, respectively. That such functor categories fit into the algebraic set theory framework was already observed in the original approach [18]. We will discuss this in detail how well-known results from topos theory translate to our framework. In particular, we will provide an explicit descriptions for the powerclass functors for the category of large sheaves.

3.1 The Category of the True Classes

We would like to think informally about the *category of the true classes* as the category having classes as objects and possibly class-sized functions between them as arrows. We will denote this category by Classes. A notion of small maps on Classes is provided by declaring a map $f : B \to A$ as small, if for all $a \in A$ the preimage $f^{-1}(a)$, is a set as opposed to be a proper class. One idea to make this rather informal definition more precise was already suggested in [18]. Let us assume that V is a universe of set theory containing an inaccessible cardinal κ. Then the elements of the universe V can be though of as the classes, whereas elements of V with cardinality strictly less than the inaccessible κ are defined to be small objects.

We will briefly indicate that the category Classes is a Heyting pretopos. Limits are calculated as in sets. The same applies to finite sums therefore they are disjoint and stable. The fact that equivalence relations have quotients is however not entirely obvious. This is due to the fact that equivalence classes can be proper classes. However, we can apply Scott's trick in order to get sets as representatives for the equivalence classes. The class containing the representing sets is easily seen to be the quotient of the equivalence relation in question.

Given a map $f : Y \to X$, the pullback functor $f^* : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ has a right adjoint $\forall_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ that maps $B \subseteq Y$ to

$$\forall_f(B) := \{x \in X \mid \text{for all } y \in f^{-1}(x), y \in B\}.$$

The natural numbers object is given by the usual natural number and is therefore stable. As announced, the small maps are the fibrewise small maps. It is straightforward to check that the axioms or the small maps are satisfied. Note, that in order to show the fibrewise small maps are closed under composition, we use the axiom of replacement. Axiom $(P2)$ follows from the powerset axiom.

3.2 Large Presheaves

Given a small category $\mathcal{C}$, presheaves over $\mathcal{C}$ are contravariant functors from $\mathcal{C}$ to the category of sets. In this section, we will allow these functors to have classes in their codomain and we will call them accordingly large sheaves. We will discuss how the category of large presheaves gives rise to a category of classes, where smallness is defines by being small componentwise. It should be remarked that the results presented in this section are standard in the topos theory literature. Our task is to check in detail that they also fit in our enlarged framework.

Definition 3.2.1. For a small category $\mathcal{C}$ the category $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$, the category of large presheaves, consists of the following:

- The objects are large presheaves, i.e. contravariant functors $X : \mathcal{C}^{op} \to$ Classes
- The arrows are natural transformations between these functors.

Proposition 3.2.2. *For any category $\mathcal{C}$ the category $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ is a Heyting pretopos with a natural number object.*

Proof.

- Limits in the category of presheaves are constructed pointwise. More precisely this means the following. Assume $F : I \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ is a diagram. Then for all $C \in \mathcal{C}$ there is a diagram

  $$F_C : I \to \text{Classes} \text{ defined by } F_C(i) := F(i)(C) \text{ and } F_C(\alpha) := F(\alpha)_C.$$

  Using that Classes is complete, for each $C \in \mathcal{C}$ there is a limiting cone X_C for F_C, in particular we have natural transformations $\rho_C : \Delta_{X_C} \Rightarrow F_C$. We will now show that there is a unique presheaf structure making the collection $\{X_C \mid C \in \mathcal{C}\}$ into a vertex of the limit for F. Observe that a given arrow $f : C' \to C$ gives rise to a cone of the diagram F'_C with vertex X_C and components $F(i)(f) \circ (\rho_C)_i$. Since X'_C is limiting, there is a unique arrow $\bar{f} : X_C \to X_{C'}$ making all involved triangles commute. So we can define a large presheaf $X : \mathcal{C}^{op} \longrightarrow$ Classes by

  $$C \mapsto X_C, \quad f : C \to C' \mapsto \bar{f} : X_C \to X_{C'}.$$

  By uniqueness of the map $\bar{f}$, X is well-defined and functorial, i.e. a large presheaf. The natural transformation $\eta : \Delta_X \Rightarrow F$ with components

  $$\eta_i : X \Rightarrow F(i) \text{ given by } (\eta_i)_C := \rho_C(i)$$

  shows that X is the vertex of a cone for F. One can easily check that X is limiting.

- Colimits in $\text{Classes}^{\mathcal{C}^{\text{op}}}$ are also computed pointwise. This can be proved similarly as above. Again, we have to pay attention that we need to use Scott's trick in order to avoid the use of the axiom of choice. So in particular finite sums are constructed pointwise. Stability under pullback and disjointness follows from the corresponding fact in Classes. Since colimits and limits are calculated pointwise in $\text{Classes}^{\mathcal{C}^{\text{op}}}$ it follows that equivalence relations have quotients. Also the stability follows from the corresponding fact in the category Classes.

- In order to show that $\text{Classes}^{\mathcal{C}^{\text{op}}}$ is a Heyting pretopos, it remains to show that for an arrow $\eta : Y \to X$, the pullback functor $\eta^* : \text{Sub}(X) \to \text{Sub}(Y)$ has a right adjoint. For $B \in \text{Sub}(Y)$ and $C \in \mathcal{C}$ define $\forall_\eta(B)(C)$ by

  $$\{x \in X(C) \mid \forall f : D \to C \text{ of } \mathcal{C}, \forall y \in Y(D)\, (\eta_D(y) = X(f)(x) \Rightarrow y \in B(D))\}$$

  Observe that for a presheaf X, the subobjects are given by the subpresheaves of X, therefore for an arrow $f : D \to C$, $\forall_\eta(B)(f)$ is just the restriction of $X(f)$ to $\forall_\eta(B)(C)$. It is easily seen that this assignment defines a functor $\forall_\eta : \text{Sub}(Y) \to \text{Sub}(X)$. Using that pullbacks in $\text{Classes}^{\mathcal{C}^{\text{op}}}$ are defined pointwise, one can easily see that for all $A \in \text{Sub}(X)$ and $B \in \text{Sub}(Y)$

  $$\eta^*(A) \leq B \quad \Leftrightarrow \quad A \leq \forall_\eta(B),$$

  so $\forall_\eta$ is indeed the right adjoint of η^*. This finishes the proof that $\text{Classes}^{\mathcal{C}^{\text{op}}}$ is a Heyting pretopos. The natural number object is given by the constant presheaf with value $\mathbb{N}$.

□

We are now going to describe the collection of small maps for the category $\text{Classes}^{\mathcal{C}^{\text{op}}}$. We want to call a natural transformation $\eta : X \Rightarrow Y$ small if all its components are small in the category of classes, so we define

$$\mathcal{S} := \{\eta : X \Rightarrow Y \mid \text{ such that for all } C \in \mathcal{C}, \eta_C : X(C) \to Y(C) \text{ is small in Classes}\}.$$

Observe that according to this definition, a large presheaf $X : \mathcal{C}^{op} \to \text{Classes}$ is small, if $X(C)$ is a set for all $C \in \mathcal{C}$.

Before we check that the axioms from Definition 2.2.1 are satisfied, we will give a precise description of the $\mathcal{P}_\text{s}$ functor.

For a large presheaf $X : \mathcal{C}^{op} \to \text{Classes}$, we define $\mathcal{P}_\text{s}(X) : \mathcal{C}^{op} \to \text{Classes}$ by

$$\mathcal{P}_\text{s}(X)(C) = \{A \mid A \text{ is a small subpresheaf of } y_C \times X\}$$

and for an arrow $f : C' \to C$ we define

$$\mathcal{P}_\text{s}(X)(f) = (y_f \times \text{id}_X)^*,$$

where y_f results from the Yoneda-embedding. So a small subpresheaf A of $y_C \times X$ will be mapped to the pullback along $y_f \times \text{id}_X$. By the pointwise description of pullbacks one can easily see that $\mathcal{P}_\text{s}(X)(f) = (y_f \times \text{id}_X)^*$ indeed maps small subpresheaves of $y_C \times X$ to small subpresheaves of $y_{C'} \times X$.

Moreover, we define $\in_\text{X} \subseteq X \times \mathcal{P}_\text{s}(X)$ by

$$\in_\text{X}(C) := \{(x, A) \in X(C) \times \mathcal{P}_\text{s}(X)(C) \mid (\text{id}_C, x) \in A(C)\}$$

Claim. *The assignment $\in_{\mathrm{X}}(C)$ gives rise to a small relation on $X \times \mathcal{P}_{\mathrm{s}}(X)$.*

Proof. First, we will show that $\in_{\mathrm{X}}$ defines a subpresheaf on $X \times \mathcal{P}_{\mathrm{s}}(X)$. To this end we need to show that for any $f : D \to C$ of $\mathcal{C}$

$$(x, A) \in \in_{\mathrm{X}}(C) \text{ implies } (X(f)(x), \mathcal{P}_{\mathrm{s}}(X)(f)(A)) \in \in_{\mathrm{X}}(D).$$

So assume $(x, A) \in \in_{\mathrm{X}}(C)$, i.e. $(\mathrm{id}_C, x) \in A(C)$. We want to show that

$$(X(f)(x), \mathcal{P}_{\mathrm{s}}(X)(f)(A)) \in \in_{\mathrm{X}}(D). \tag{3.1}$$

To this end we would need to check that $(\mathrm{id}_D, X(f)(x)) \in \mathcal{P}_{\mathrm{s}}(X)(f)(A)(D)$. Since $\mathcal{P}_{\mathrm{s}}(X)(f)$ is defined by pullback along $y_f \times \mathrm{id}_X$ and the pullback is defined pointwise, the latter holds precisely when $(f, X(f)(x)) \in A(D)$. Now since A itself is a subpresheaf and by assumption $(\mathrm{id}_C, x) \in A(C)$ it follows that $(y_C \times X)(f)(\mathrm{id}_C, x) \in A(D)$. But since $(y_C \times X)(f)(\mathrm{id}_C, x) = (f, X(f)(x))$, we have established (3.1). Therefore we showed that $\in_{\mathrm{X}}$ defines a subpresheaf of $X \times \mathcal{P}_{\mathrm{s}}(X)$. It remains to show that this is a small relation. So we need to show that the composite

$$\in_{\mathrm{X}} \rightarrowtail X \times \mathcal{P}_{\mathrm{s}}(X) \xrightarrow{\pi_2} \mathcal{P}_{\mathrm{s}}(X)$$

defines a small map. Since small maps are defined pointwise, this reduces to show that for every $C \in \mathcal{C}$ the composite

$$\in_{\mathrm{X}}(C) \rightarrowtail X(C) \times \mathcal{P}_{\mathrm{s}}(X)(C) \xrightarrow{\pi_2} \mathcal{P}_{\mathrm{s}}(X)(C)$$

is a small map in Classes. So assume $A \in \mathcal{P}_{\mathrm{s}}(X)(C)$, i.e. A is a small subpresheaf of $y_C \times X$. In order to show that the above composition is small in Classes, we need to show that there are just set-many $x \in X(C)$ with $(x, A) \in \in_{\mathrm{X}}(C)$, i.e. $(\mathrm{id}_C, x) \in A(C)$. But since A is small and therefore $A(C)$ is a set for every $C \in \mathcal{C}$, the latter follows. $\square$

Proposition 3.2.3. *The class $\mathcal{S}$ defines a class of small maps.*

Proof. We will check that the conditions from Definition 2.2.1 are satisfied. Since smallness, limits and colimits are defined pointwise, it immediately follows that (S1)-(S5) are satisfied. We check that the representability axiom (P1) is validated.

Above, we already argued that $\in_{\mathrm{X}}$ defines a small relation on $X \times \mathcal{P}_{\mathrm{s}}(X)$. It remains to show the universal property of this relation. So assume $R \subseteq X \times Y$ defines a small relation on $X \times Y$. Define the map $t : Y \Rightarrow \mathcal{P}_{\mathrm{s}}(X)$ by components

$$\begin{aligned} t_C : Y(C) &\longrightarrow \mathcal{P}_{\mathrm{s}}(X)(C) \\ y &\mapsto t_C(y), \end{aligned}$$

where $t_C(y) : \mathcal{C}^{op} \longrightarrow \text{Classes}$ is the subpresheaf of $y_C \times X$ given by

$$t_C(y)(D) := \{(g, x) \mid g : D \to C, x \in X(D), \text{ such that } (Y(g)(y), x) \in R(D)\}.$$

That this assignment gives rise to a subpresheaf of $y_C \times X$, follows from R being a subpresheaf of $X \times Y$. Or more precisely, assume $(g, x) \in t_C(y)(D)$ and let f :

$D' \to D$ be a map in $\mathcal{C}$. In order to show that $(y_C \times X)(f)(g, x) \in t_C(y)(D')$, we need to check that $(Y(f)(Y(g)(y)), X(f)(x)) = (Y(gf)(y), X(f)(x)) \in R(D')$. But the latter follows by R being a subpresheaf. Let us turn to the proof that $t_C(y)$ is small. For this purpose, we need to show that for every $D \in \mathcal{C}$, $t_C(y)(D)$ is a set. Since R is small, for every $g : D \to C$ there are just set-many elements $x \in X(D)$ with $(Y(g)(y), x) \in R(D)$. Moreover, since $\mathcal{C}$ is small, there are also just set many candidates for the first component of an element in $t_C(y)(D)$, therefore the latter needs to be a set. One can check that $t : Y \to \mathcal{P}_{\mathrm{s}}(X)$ is the unique map fitting in the required pullback diagram.

In order to show the powerset axiom (P2), observe that for a presheaf X, the internal subset relation $\subseteq_X \colon \mathcal{C}^{\mathrm{op}} \to$ Classes is given by components

$$\subseteq_X (C) = \{(A, B) \in \mathcal{P}_{\mathrm{s}}(X)(C) \times \mathcal{P}_{\mathrm{s}}(X)(C) \mid A \text{ is a subpresheaf of } B\}.$$

It is clear that the above defines a small relation. If $B \in \mathcal{P}_{\mathrm{s}}(X)(C)$, then B is a small subpresheaf of $y_C \times X$. Using the powerset axiom in Classes, it is clear that there are only set-many subpresheaves of B. Smallness of the natural number object follows since it is pointwise small. $\square$

By our considerations in Chapter 2, we know that the assignment

$$X \mapsto \mathcal{P}_{\mathrm{s}}(X)$$

is in fact functorial. The action of $\mathcal{P}_{\mathrm{s}}$ on arrows in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ can also easily be described explicitly. If $\eta : X \Rightarrow Y$ is and arrow in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$, then $\mathcal{P}_{\mathrm{s}}(\eta)_C : \mathcal{P}_{\mathrm{s}}(X)(C) \to \mathcal{P}_{\mathrm{s}}(Y)(C)$ maps a small subpresheaf $Z \subseteq y_C \times X$ to the image of Z under $\mathrm{id}_{y_C} \times \eta$. So for some C' in $\mathcal{C}$ we have

$$\begin{aligned}\mathcal{P}_{\mathrm{s}}(\eta)_C(Z)(C') :=& \{(h, y) \in y_C(C') \times Y(C') \mid \exists x \in X(C') \text{ such that} \\ & (h, x) \in Z(C') \text{ and } \eta_{C'}(x) = y\}.\end{aligned}$$

Remark 3.2.4. Note that in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$, we have in fact class-sized colimits. Analogously to the well-known fact for presheaves, see for example Proposition 1.1. in [23], one can show that every large presheaf X is a colimit of representables over a (possibly class sized) diagram. This property implies that the representable presheaves form a system of generators for $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$, i.e. for any two distinct arrows $f \neq g : X \Rightarrow Y$ between large presheaves X and Y, there is a representable presheaf y_C and a map $h : y_C \Rightarrow X$ such that $fh \neq gh$.

3.3 Large Sheaves on a Site

The last example that we examine is the category of large sheaves over a small category $\mathcal{C}$. Analogously to the case a presheaves, large sheaves are allowed to have proper classes in their codomain. In order to show that the large sheaves on a site give rise to a category of classes, we will develop the most basic results from sheaf theory including the notion of the associated sheaf functor. Again, the presented results are standard, we will only check that their "enlarged version" still holds. At the end of this section, we will have a closer look at the dense topology and examine large sheaves over poset categories.

3.3.1 Grothendieck Topologies

Grothendieck topologies generalize the notion of an open cover on a topological space. A category $\mathcal{C}$ together with Grothendieck topology Cov is called a *site*. We will see how this additional structure gives rise to the notion of large sheaves. Let $\mathcal{C}$ be a small category and C be an object of $\mathcal{C}$. A *sieve* R on C is a subpresheaf $R \subseteq y_C$ of the representable presheaf y_C. Equivalently, R is a collection of arrows of $\mathcal{C}$ with codomain C that is in a sense downwards closed, meaning that if $h : D \to C$ is in R, then for every arrow $g : B \to D$ the composite $h \circ g$ is an element of R. Now assume $f : C' \to C$ is an arrow in $\mathcal{C}$ and R is a sieve on C. Then it is easy to see that the set

$$f^*(R) := \{g : B \to C' \mid f \circ g \in R\}$$

defines a sieve on C'. Note that in fact $f^*(R) = {y_f}^*(R)$, where ${y_f}^*$ is the pullback functor associated to the natural transformation $y_f : y_{C'} \Rightarrow y_C$ resulting from the Yoneda embedding. A Grothendieck topology assigns to every object C of $\mathcal{C}$ a collection of covering sieves.

Definition 3.3.1. Let $\mathcal{C}$ be a small category. A *Grothendieck topology* on $\mathcal{C}$ assigns to every $C \in \mathrm{Ob}(\mathcal{C})$ a set $\mathrm{Cov}(C)$ of sieves on C, such that the following conditions are satisfied:

- The maximal sieve on C is in $\mathrm{Cov}(C)$.
- If $R \in \mathrm{Cov}(C)$ and $f : C' \to C$ is an arrow in $\mathcal{C}$, then $f^*(R) \in \mathrm{Cov}(C')$.
- If $R \in \mathrm{Cov}(C)$ and S is a sieve on C such that for every $f : C' \to C$ in R we have that $f^*(S) \in \mathrm{Cov}(C')$, then $S \in \mathrm{Cov}(C)$.

One can easily verify the following properties.

Lemma 3.3.2. *For every $C \in \mathrm{Ob}(\mathcal{C})$*

- *If $R \in \mathrm{Cov}(C)$ and S a sieve on C with $R \subseteq S$, then $S \in \mathrm{Cov}(C)$.*
- *If $R, S \in \mathrm{Cov}(C)$, then $R \cap S \in \mathrm{Cov}(C)$.*

If $R \in \mathrm{Cov}(C)$, we call R a *cover* for C or a *covering sieve*. A pair $(\mathcal{C}, \mathrm{Cov})$ of a small category $\mathcal{C}$ and a Grothendieck topology Cov on $\mathcal{C}$ is called a *site*.

Definition 3.3.3. Let $\mathcal{C}$ be a small category, $C \in \mathrm{Ob}(\mathcal{C})$ and R a sieve on C. Let $X : \mathcal{C}^{\mathrm{op}} \to \mathrm{Classes}$ be a large presheaf. A *compatible family* for X indexed by R is given by elements $x_f \in X(C')$ for all $f : C' \to C \in R$, such that for all $g : C'' \to C' \in \mathrm{Ar}(\mathcal{C})$

$$x_{fg} = X(g)(x_f).$$

Since a sieve R on C is the same as a subpresheaf of the representable functor y_C, an alternative description of a compatible family for X indexed by R is a natural transformation

$$\eta : R \Rightarrow X.$$

Naturality of η corresponds precisely to compatibility of the family $(\eta_{C'}(f))_{f:C'\to C\in R}$. If $(x_f)_{f\in R}$ is a compatible family for X indexed by R, an *amalgamation* for the family is an element $x \in X(C)$ such that

$$x_f = X(f)(x) \text{ for all } f \in R.$$

Definition 3.3.4. Let $(\mathcal{C}, \mathrm{Cov})$ be a site. A large presheaf $X : \mathcal{C}^{op} \to \mathrm{Classes}$ is called a *large sheaf*, if every compatible family on X indexed by some covering sieve has exactly one amalgamation.

There is also a weaker notion, namely the one of a separated large presheaf. Such a presheaf satisfies the uniqueness condition in the above definition but not necessarily the existence condition.

Definition 3.3.5. A large presheaf X is called *separated*, if every compatible family on X has at most one amalgamation.

The large sheaves on a site determine a full subcategory of the large presheaves.

Definition 3.3.6. Given a site $(\mathcal{C}, \mathrm{Cov})$, we denote by $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$, the full subcategory of $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ generated by the sheaves for $(\mathcal{C}, \mathrm{Cov})$, i.e.

- The objects of $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ are the large sheaves for $(\mathcal{C}, \mathrm{Cov})$.
- For $X, Y \in \mathrm{Ob}(\mathrm{LSh}(\mathcal{C}, \mathrm{Cov}))$, the arrows between X and Y are all arrows in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ between X and Y regarded as large presheaves.

The goal of the following sections is to prove that the category of large sheaves on a site forms a category of classes. Before we can prove this fact, it will be useful to establish the notion of the associated sheaf functor.

3.3.2 The Associated Sheaf Functor

In the present section we will always work with a fixed site $(\mathcal{C}, \mathrm{Cov})$. As the large sheaves on $(\mathcal{C}, \mathrm{Cov})$ form a subcategory of the large presheaves, there is an inclusion functor

$$\iota : \mathrm{LSh}(\mathcal{C}, \mathrm{Cov}) \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} .$$

The goal of this section is to show that ι has a left adjoint

$$\mathbf{a} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{LSh}(\mathcal{C}, \mathrm{Cov}) .$$

The functor $\mathbf{a}$ is called the *associated sheaf functor* or *sheafification functor*. The construction of $\mathbf{a}$ proceeds in two steps. In a first step one constructs for a large presheaf $X : \mathcal{C}^{op} \to \mathrm{Classes}$ a large presheaf $X^+ : \mathcal{C}^{op} \to \mathrm{Classes}$. This process is referred to as the $^+$-construction. The presheaf X^+ is not necessarily a sheaf, but it is already separated. Since the $^+$-construction always turns a separated large presheaf into a sheaf, the result of applying the $^+$-construction twice will always be sheaf.

Let us describe the $^+$-construction. Let $X : \mathcal{C}^{op} \to \mathrm{Classes}$ be a large presheaf. We define a relation on pairs of the form (R, ϕ), where $R \in \mathrm{Cov}(C)$ and $\phi : R \Rightarrow X$ an arrow in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ by

$$(R, \phi) \sim (S, \psi) \Leftrightarrow \exists T \in \mathrm{Cov}(C) \text{ with } T \subseteq R \cap S, \text{ such that } \phi \text{ and } \psi \text{ coincide on } T.$$

Note that every pair (R, ϕ) stands for a compatible family on X indexed by R. Now $\sim$ is the relation that identifies two such compatible families, if they agree on a covering sieve. The relation $\sim$ is easily seen to be an equivalence relation by using the properties of covering sieves that we stated in Lemma 3.3.2. Note, that since $\mathcal{C}$ is small, all natural transformations $\phi : R \Rightarrow X$ are sets. However, the equivalence classes of $\sim$ might be proper classes. We can apply Scott's trick to choose sets of representatives for the equivalence classes. We denote the (set of representatives of an) equivalence class of a pair (R, ϕ) by $[(R, \phi)]$. We define $X^+ : \mathcal{C}^{op} \to \text{Classes}$ on objects $C \in \text{Ob}(\mathcal{C})$ by

$$X^+(C) = \{[(R, \phi)] \mid R \in \text{Cov}(C), \phi : R \Rightarrow X \text{ nat. trans.}\}.$$

And for an arrow $f : C' \to C \in \text{Ar}(\mathcal{C})$, we define

$$\mathrm{X}^+(f) : \mathrm{X}^+(C) \to \mathrm{X}^+(C'), [(R, \phi)] \mapsto [(f^*(R), \psi)],$$

where $\psi = \phi \circ f'$ and f' is as in the pullback

$$\begin{tikzcd} f^*(R) \arrow[r, "f'"] \arrow[d] & R \arrow[d] \\ y_C \arrow[r, "y_f"'] & y_{C'} \end{tikzcd}$$

It is easy to see that $X^+(f)$ is well-defined on equivalence classes. Also the functorality of X^+ is easy to verify.

The assignment $X \mapsto X^+$ extends to a functor

$$(-)^+ : \text{Classes}^{\mathcal{C}^{\text{op}}} \to \text{Classes}^{\mathcal{C}^{\text{op}}},$$

by setting for $\eta : X \Rightarrow Y \in \text{Ar}(\text{Classes}^{\mathcal{C}^{\text{op}}})$,

$$(\eta^+)_C : X^+(C) \Rightarrow Y^+(C), [(R, \phi)] \mapsto [(R, \phi \circ \eta)].$$

Again, one can easily check that this is well-defined on equivalence classes and the functorality of the assignment.

As in Lemma 2.11 and 2.12 in [23], one can prove that the $^+$-constructions always results in a separated presheaf and that it turns separated presheaves into sheaves.

Lemma 3.3.7. *For every large presheaf X, the large presheaf X^+ is separated.*

Lemma 3.3.8. *If X is separated, then X^+ is a large sheaf.*

Lemma 3.3.7 and 3.3.8 show that by applying the $^+$-construction to a large presheaf twice, the result will always be a large sheaf. We define the associated sheaf functor by

$$\begin{aligned} \mathbf{a} : \text{Classes}^{\mathcal{C}^{\text{op}}} &\to \text{LSh}(\mathcal{C}, \text{Cov}) \\ X &\mapsto X^{++} \\ \eta : X \Rightarrow Y &\mapsto \eta^{++}. \end{aligned}$$

In order to show that $\mathbf{a}$ is indeed left adjoint to the inclusion, we define a natural transformation

$$\zeta : \text{id}_{\text{Classes}^{\mathcal{C}^{\text{op}}}} \Longrightarrow (-)^+$$

with components

$$\begin{aligned}\zeta_X : X &\Longrightarrow \mathrm{X}^+ \text{ where, for } C \text{ in } \mathcal{C}\\ (\zeta_X)_C : X(C) &\to \mathrm{X}^+(C)\\ x &\mapsto [(y_C, \phi)],\end{aligned}$$

where $\phi : y_C \Rightarrow X$ corresponds to $x \in X(C)$ according to the Yoneda Lemma. The natural transformation ζ has a universal property as proved for example in 2.10 in [23].

Proposition 3.3.9. *Let X be a large presheaf, F a large sheaf and $g : X \Rightarrow F$ an arrow in* $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$*. Then g factors uniquely through $\zeta_X : X \to X^+$ as in*

$$\begin{tikzcd} X \arrow[rr, "g"] \arrow[dr, "\zeta_X"'] & & F \\ & X^+ \arrow[ur, dashed, "\tilde{g}"'] & \end{tikzcd}.$$

Using the universal property stated above, it is easy to check:

Corollary 3.3.10. *Assume $\zeta_X : X \to X^+$ factors through a large presheaf X', X' separated, then*

$$(X')^+ \cong X^{++}.$$

Now define $\eta : \mathrm{id}_{\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}} \Longrightarrow \mathbf{a}$ by $\eta := \zeta \circ \zeta$. Again by applying Proposition 3.3.9 twice, it immediately follows that η serves as the unit of the desired adjunction. So we established:

Proposition 3.3.11. *The functor* $\mathbf{a} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ *is left adjoint to the inclusion functor.*

Being left adjoint, the sheafification functor preserves colimits. However, one can also show that it also preserves finite limits. As in Lemma 2.13 in [23], the $^+$-construction preserves finite limits. Again, by applying the result twice, we obtain:

Corollary 3.3.12. *The associated sheaf functor* $\mathbf{a} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ *preserves finite limits.*

3.3.3 Closed Subpresheaves

In this section we will show that a site $(\mathrm{Cov}\,\mathcal{C})$ determines a universal closure operation on the category $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$, i.e. a closure operation on all subobject lattices $\mathrm{Sub}(X)$ for $X \in \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ that is stable under pullback. The closed subpresheaves of a large presheaf are the ones that remain invariant under this closure operation. Given a presheaf X, a closed subpresheaf of X contains amalgamations for all compatible families on it. If X itself is a sheaf, then its closed subpresheaves are exactly its subsheaves. We will encounter the closed subpresheaves again, when we define the $\mathcal{P}_{\mathrm{s}}$-functor for $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ is the next section. We begin with the definition of a universal closure operation.

Definition 3.3.13. A *universal closure operation* on $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ assigns to every large presheaf X an operation $\overline{(\cdot)} : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ such that the following hold

(1) For all $A \in \mathrm{Sub}(X)$ we have $A \leq \bar{A}$ and

(2) $\bar{A} = \bar{\bar{A}}$.

(3) If $A \leq B$ in $\mathrm{Sub}(X)$ then $\bar{A} \leq \bar{B}$.

(4) For all $\phi : Y \Rightarrow X$ and for all $A \in \mathrm{Sub}(X)$, $\phi^*(\bar{A}) = \overline{\phi^*(A)}$, where $\phi^* : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ is the associated pullback functor.

Properties (1)-(3) state that for every $X \in \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ the function $\overline{(\cdot)}$: $\mathrm{Sub}(X) \to \mathrm{Sub}(X)$ is indeed a closure operation in the usual sense. Property (4) requires that the various closure operations are connected in the sense that closure is stable under pullback. The next proposition asserts that every site determines a universal closure operation.

Proposition 3.3.14. *Let* $(\mathcal{C}, \mathrm{Cov})$ *be a site. Then* $(\mathcal{C}, \mathrm{Cov})$ *determines a universal closure operation on* $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$.

Proof. Let $X \in \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ and $Z \in \mathrm{Sub}(X)$. We define for $C \in \mathrm{Ob}(\mathcal{C})$, and $x \in X(C)$

$$x \in \overline{Z}(C) \quad \Leftrightarrow \quad \begin{array}{l} \text{There is a covering sieve } R \text{ on } C \text{ and} \\ X(f)(x) \in Z(C') \text{ for all } f : C' \to C \text{ in } R. \end{array}$$

By using the properties of covering sieves that we established earlier, it is easily checked that the above definition determines a universal closure operation as desired. □

The notion of a closure operation gives rise to the definition of closed subpresheaves of a given presheaf. Note that the definition of closure very much depends on the Grothendieck topology that we are working with.

Definition 3.3.15. Let $(\mathcal{C}, \mathrm{Cov})$ be a site and $X \in \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$. A subpresheaf Z of X is called *closed* if $\bar{Z} = Z$ in $\mathrm{Sub}(X)$, where $\overline{(\cdot)} : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ is part of the universal closure operation that $(\mathcal{C}, \mathrm{Cov})$ determines.

By the definition of the closure operator on a large presheaf X, the proof of the following lemma is immediate.

Lemma 3.3.16. *A subpresheaf* Z *of* X *is closed if and only if for all* $x \in X(C)$*: If* $\{f : C' \to C \mid X(f)(x) \in Z(C')\}$ *covers* C*, then* $x \in Z(C)$.

As announced, it turns out that the closed subpresheaves of a large sheaf are precisely the ones that are sheaves themselves.

Lemma 3.3.17. *Let* X *be a sheaf. Then a subpresheaf* $Z \subseteq X$ *is a sheaf if and only if* Z *is closed.*

Proof. For the direction from right to left, assume that $Z \in \mathrm{Sub}(X)$ is closed. It is easy to see that subpresheaves of sheaves are separated. Therefore an existing amalgamation of a compatible family for Z is necessarily unique. The existence of amalgamations is easily seen by Lemma 3.3.16. The direction from left to right is clear by the definition of the closure operation in Proposition 3.3.14. □

This finishes our considerations on closed subpresheaves. It should be remarked that in fact the converse of Proposition 3.3.14 is also true: every universal closure operation on a category of large presheaves $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ determines a Grothendieck topology on the index category $\mathcal{C}$ and the notions are interdefinable. A proof of this fact is worked out in Theorem 2.5 of [23].

3.3.4 Large Sheaves as a Category of Classes

We will now show that the large sheaves on a site give rise to a category of classes. The small maps are inherited from the presheaves. The pointwise small maps give rise to small maps on $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$.

Proposition 3.3.18. $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ *is closed under limits in* $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$.

Proof. Let $\mathcal{F} : I \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ be a diagram such that for all $i \in I$, $\mathcal{F}(i)$ is a large sheaf. Let $X : \mathcal{C}^{op} \to \mathrm{Classes}$ be the limiting cone for $\mathcal{F}$ in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$.

Claim. X *is a sheaf.*

In order to prove the claim, let $R \in \mathrm{Cov}(C)$ be a covering sieve on C and let $\nu : R \Rightarrow X$ be a natural transformation. We will show that there is a unique $x \in X(C)$, with $\nu_{C'}(f) = X(f)(x)$ for all $f : C' \to C \in R$. The proof will use the pointwise construction of limits for presheaves and the sheaf structure of the $\mathcal{F}(i)$s. Since X is a cone for $\mathcal{F}$, there is a natural transformation $\eta : \Delta_X \Rightarrow \mathcal{F}$ with components $\eta_i : X \Rightarrow \mathcal{F}(i)$ for all $i \in I$. Then for all $i \in I$, the composition

$$\eta_i \circ \nu : R \Rightarrow \mathcal{F}(i)$$

defines a compatible family for $\mathcal{F}(i)$ indexed by R. Since $\mathcal{F}(i)$ is a large sheaf, there exists $x_i \in \mathcal{F}(i)(C)$, such that for all $f : C' \to C \in R$,

$$(\eta_i \circ \nu)_{C'}(f) = \mathcal{F}(i)(f)(x_i). \tag{3.2}$$

Observe, that for all $k : i \to j$, we have $\mathcal{F}(k)_C(x_i) = x_j$. This follows since for all $f : C' \to C \in R$ we have

$$\begin{aligned} & \mathcal{F}(j)(f)(x_j) \\ = \;& (\eta_j \circ \nu)_{C'}(f) \\ = \;& (\mathcal{F}(k) \circ \eta_i \circ \nu)_{C'}(f) \\ = \;& \mathcal{F}(k)_{C'}(\mathcal{F}(i)(f)(x_i)) \\ = \;& \mathcal{F}(j)(f)(\mathcal{F}(k)_C(x_i)), \end{aligned}$$

where the first equality follows by η being natural, the second by (3.2) and the last one by naturality of $\mathcal{F}(k)$. By uniqueness of the amalgamation for the compatible family $\eta_j \circ \nu$, it follows that $x_j = \mathcal{F}(k)_C(x_i)$. This shows that the terminal object $\{*\}$ is a cone for the diagram $\mathcal{F}(-)(C) : I \to \mathrm{Classes}$ with components given by $* \mapsto x_i$. Since limits in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ are constructed pointwise, $X(C)$ is the limiting cone of $\mathcal{F}(-)(C) : I \to \mathrm{Classes}$. Therefore, there is a unique element $x \in X(C)$ with

$$(\eta_i)_C(x) = x_i \tag{3.3}$$

We have for all $f : C' \to C \in R$,

$$\begin{array}{rl} & (\eta_i \circ \nu)_{C'}(f) \\ = & \mathcal{F}(i)(f)\left((\eta_i)_C(x)\right) \\ = & (\eta_i)_{C'}(X(f)(x)), \end{array}$$

where the first equality follows by (3.2) and (3.3) and the second by naturality of η_i. By the universal property of η_i it follows that $X(f)(x) = \nu_{C'}(f)(x)$. This shows that our $x \in X(C)$ is the unique amalgamation for the compatible family defined by ν. This proves that X is a sheaf. $\square$

Proposition 3.3.19. $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ *is a Heyting pretopos with a stable natural number object.*

Proof. First, we will prove that $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ is a Heyting pretops.

- Finite limits exists by Proposition 3.3.18.
- Given a diagram $\mathcal{F} : I \Rightarrow \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$, the colimit of this diagram is the sheaf X^{++}, where X is the limiting cocone in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$. This follows as **a** is left adjoint to the inclusion. So $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ is cocomplete. For sums we also have a more direct description. For $X, Y \in \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$, define $(X+Y)' : \mathcal{C}^{op} \to \mathrm{Classes}$ by

  $$(X+Y)'(C) = \left\{ \begin{array}{ll} X + Y(C) & \text{if } \emptyset \notin \mathrm{Cov}(C) \\ \{*\} & \text{if } \emptyset \in \mathrm{Cov}(C) \end{array} \right\}$$

  where $+$ denotes the sum of A and B in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$. One can show that $(X+Y)'$ is separated and that $X+Y \to (X+Y)^+$ factors through $(X+Y)'$ and therefore by Corollary 3.3.10

  $$((X+Y)')^+ \cong (X+Y)^{++}.$$

  We will now sketch the proof that coproducts are stable under pullback. Let $f : Y \Rightarrow X$ be an arrows in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$. And assume that A and B are subsheaves of X. We want to show that $f^*\left(((A+B)')^+\right) \cong \left((f^*(A) + f^*(B))'\right)^+$, where again $+$ denotes the sum of A and B in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$. We have

  $$\begin{array}{rl} & f^*\left(((A+B)')^+\right) \\ \cong & \left(f^*\left((A+B)'\right)\right)^+ \\ \cong & \left((f^*(A) + f^*(B))'\right)^+, \end{array}$$

  where the first $\cong$ follows since $(-)^+$ preserves finite limits by our considerations in Section 3.3.2 and the second since sums are stable in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ and the definition of $(-)'$. Disjointness follows from the corresponding fact for large presheaves.
- That equivalence relations have quotients that are stable follows from the corresponding fact in presheaves using that the sheafification functor preserves finite limits and coequalizers.

  We just argued that $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ is cocomplete. Therefore, equivalence relations have quotients that are easily checked to be stable.

- Finally, we will show that pullback functors have right adjoints. So let $\eta : Y \Rightarrow X$ be an arrow in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$. Note that by Proposition 3.3.18, the pullback functor $\eta^* : \mathrm{Sub}(X) \Rightarrow \mathrm{Sub}(Y)$ is the restriction of the pullback functor corresponding to η in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ to the subpresheaves of X that are sheaves. Accordingly, we can try to define $\forall_\eta : \mathrm{Sub}(Y) \Rightarrow \mathrm{Sub}(X)$ as in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ by setting for a subsheaf $B \in \mathrm{Sub}(Y)$

$$\begin{aligned}\forall_\eta(B)(C) = \{x \in X(C) \mid \forall f : D \to C \text{ of } \mathcal{C}, \forall y \in Y(D) \\ (\eta_D(y) = X(f)(x) \Rightarrow y \in B(D))\}.\end{aligned}$$

  We already proved that this defines a subpresheaf of X. We will show that $\forall_\eta(B)$ is actually also a sheaf. So assume R is a covering sieve on $C \in \mathcal{C}$ and $\nu : R \Rightarrow \forall_\eta(B)$ a natural transformation. As a subpresheaf of a sheaf, $\forall_\eta(B)$ is separated. So it remains to show that there is an amalgamation of the compatible family determined by ν. It is clear that ν also defines a compatible family on X and since X is a sheaf, there is $x \in X(C)$ such that

$$\text{for all } f : C' \to C \in R;\ \nu_{C'}(f) = X(f)(x) \tag{3.4}$$

  We need to show that in fact, $x \in \forall_\eta(B)$. To this end, let $g : D \to C \in \mathcal{C}$ and $y \in Y(C)$ with

$$\eta_D(y) = X(g)(x). \tag{3.5}$$

  We need to show that $y \in B(D)$. Observe that $g^*(R)$ is a covering sieve on D. For $h : E \to D \in g^*(R)$, $Y(h)(y)$ determines a compatible family for Y, indexed by the covering sieve $g^*(R)$ whose unique amalgamation is y. So if we are able to show that for all $h : E \to D \in g^*(R)$, $Y(h)(y) \in B(E)$ it would follow that $y \in B(D)$ since B is a sheaf. So let $h : E \to D \in g^*(R)$. Then $g \circ h \in R$ and so clearly, $\nu_E(gh) \in \forall_\eta(B)(E)$. We have

$$\begin{aligned} & \eta_E\left((Y)(h)(y)\right) & \\ = \; & X(h)\left(\eta_D(y)\right) & \text{[by naturality of } \eta] \\ = \; & X(h)\left(X(g)(x)\right) & \text{[by (3.5)]} \\ = \; & \nu_E(gh) & \text{[by (3.4)]} \\ = \; & X(id_E)(\nu_E(gh)). & \end{aligned}$$

  Since $\nu_E(gh) \in \forall_\eta(B)(E)$ this implies that $(Y)(h)(y) \in B(E)$. By our above argumentation, this finishes the proof that $\forall_\eta(B)$ is a sheaf. That $\forall_\eta$ is a right adjoint of η^* follows as in the case of large presheaves.

This finishes the proof that the large sheaves form a Heyting pretopos. The natural number object is given by the sheafification of the natural number object of the large presheaves. This finishes the proof of the proposition. □

Now we can equip $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ with a class of maps

$$\mathcal{S} := \{\eta : X \Rightarrow Y \mid \eta \text{ is a small map in } \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}.\}$$

Proposition 3.3.20. *The class $\mathcal{S}$ satisfies the axioms for small maps.*

The validity of the axioms $(S1)$, $(S2)$ and $(S4)$ are immediate as limits in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ are constructed pointwise. To check the validity of axiom $(S5)$, one uses the precise description of the coproducts given above. As we already remarked at the end of Section 2.2, Simpson [28] showed that $(S3)$ in fact is implied by the validity of the other axioms, so we do not need to check its validity separately. We will check $(P1)$. To this end we will provide a precise description of the $\mathcal{P}_{\mathrm{s}}$ -functor for large sheaves on a site. For $X \in \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ we define $\mathcal{P}_{\mathrm{s}}(X) : \mathcal{C}^{op} \to \mathrm{Classes}$ by setting for $C \in \mathrm{Ob}(\mathcal{C})$

$$\mathcal{P}_{\mathrm{s}}(X)(C) = \{Z \mid Z \text{ is a small and closed subpresheaf of } y_C \times X\},$$

and for $f : C' \to C \in \mathrm{Ar}(\mathcal{C})$ we define the action of $\mathcal{P}_{\mathrm{s}}(X)$ by pullback along $y_f \times \mathrm{id}_X$, i.e. we define

$$\mathcal{P}_{\mathrm{s}}(X)(f) = (y_f \times \mathrm{id}_X)^*.$$

By property (4) in Definition 3.3.13 it follows that $(y_f \times \mathrm{id}_X)^*$ maps a closed subpresheaf of $y_C \times X$ to a closed subpresheaf of $y_{C'} \times X$. By the pointwise description of pullbacks it is also immediate to see that $(y_f \times \mathrm{id}_X)^*$ also preserves smallness of the subpresheaves. Therefore, $\mathcal{P}_{\mathrm{s}}(X)(f)$ is well-defined.

Proposition 3.3.21. *For every large presheaf X, $\mathcal{P}_{\mathrm{s}}(X)$ is a sheaf.*

Proof. First, we show that $\mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is separated. Assume Y and Y' are small and closed subpresheaves of $y_C \times X$. Let $R \in \mathrm{Cov}(C)$ such that for all $f : C' \to C \in R$ we have

$$(y_f \times X)^*(Y) = (y_f \times X)^*(Y'). \tag{3.6}$$

We need to show that $Y = Y'$. For the latter it is enough to show that for all $D \in \mathcal{C}$, $Y(D) \subseteq Y'(D)$, since the other direction follows by symmetry. Assume $(f, x) \in Y(D)$. If $f : D \to C \in R$, we have

$$\begin{aligned} & (f, x) \in Y(D) \\ \Rightarrow \quad & (\mathrm{id}_D, x) \in (y_f \times X)^*(Y)(D) \\ \Rightarrow \quad & (\mathrm{id}_D, x) \in (y_f \times X)^*(Y')(D) \\ \Rightarrow \quad & (f, x) \in Y'(D). \end{aligned}$$

Here, the first and the last implication follow from the definition of the pullback and the one in the middle follows from (3.6). Now assume that $f : D \to C$ is arbitrary. Then $f^*(R)$ is a covering sieve on D. Let $k : E \to D \in f^*(R)$, i.e. $f \circ k \in R$.

We have

$$\begin{aligned} & (f, x) \in Y(D) \\ \Rightarrow \quad & (f \circ k, X(f)(x)) \in Y(E) \\ \Rightarrow \quad & (f \circ k, X(f)(x)) \in Y'(E), \end{aligned}$$

where the first implication follows by Y being a subpresheaf of $y_C \times X$ and the latter by $f \circ k \in R$ and the calculation above. So for every $k : E \to D \in f^*(R)$, we have $(f \circ k, X(f)(x)) = (y_C \times X)(k)(f, x) \in Y'(E)$. This means that

$$f^*(R) \subseteq \{k : E \to D \mid (y_C \times X)(k)(f, x) \in Y'(E)\} =: S.$$

Since the set of covering sieves is upwards closed and $f^*(R) \in \mathrm{Cov}(D)$, it follows that S is a covering sieve on D. Now since Y' is closed, by Lemma 3.3.16 we can conclude that $(f, x) \in Y'(D)$. This finishes the proof that $\mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is separated. It remains to show that every compatible family has an amalgamation. To this end let R be a covering sieve on C and let $(Z_f)_{f\in R}$ be a compatible family for R. Compatibility means that for every $g : E \to D \in \mathrm{Ar}(\mathcal{C})$ and $f : D \to C \in R$ we have that $Z_{fg} = (y_g \times \mathrm{id}_X)^*(Z_f)$. We need to find an amalgamation for this family. We define for $E \in \mathcal{C}$

$$Z(E) = \{(fg, x) \mid f \in R, (g, x) \in Z_f(E)\}$$

It is easy to see that Z defines a small subpresheaf of $y_C \times X$. We show that for every $f : D \to C \in R$ we have $(y_f \times \mathrm{id}_X)^*(Z) = Z_f$. The direction $Z_f \subseteq (y_f \times \mathrm{id}_X)^*(Z)$ follows directly by the definition of Z. We show that for every $E \in \mathcal{C}$ we have $(y_f \times \mathrm{id}_X)^*(Z)(E) \subseteq Z_f(E)$. Assume $(h, x) \in (y_f \times \mathrm{id}_X)^*(Z)(E)$. Then $(fh, x) \in Z(E)$. This means that there exists $f' : B \to C \in R$ and $g : E \to B$ with $(g, x) \in Z_{f'}(E)$ and $f' \circ g = f \circ h$. The latter clearly implies that $Z_{f'g} = Z_{fh}$. So by compatibility, it follows that

$$(y_g \times \mathrm{id}_X)^*(Z_{f'}) = (y_h \times \mathrm{id}_X)^*(Z_f) \tag{3.7}$$

Now we have that

$$\begin{array}{ll} & (g, x) \in Z_f(E) \\ \Rightarrow & (\mathrm{id}_E, x) \in (y_g \times \mathrm{id}_X)^*(Z_{f'})(E) \\ \Rightarrow & (\mathrm{id}_E, x) \in (y_h \times \mathrm{id}_X)^*(Z_f)(E) \\ \Rightarrow & (h, x) \in Z_f(E), \end{array}$$

where the implication in the middle follows from (3.7). Since $(h, x) \in Z_f(E)$ this shows that $(y_f \times \mathrm{id}_X)^*(Z)(E) \subseteq Z_f(E)$ and therefore we can conclude that

$$(y_f \times \mathrm{id}_X)^*(Z) = Z_f. \tag{3.8}$$

It seems that we found an amalgamation for our compatible family. Note that Z does not necessarily define a closed subpresheaf of $y_C \times X$. However, the closure $\bar{Z}$ is easily seen to be small and it is clearly a closed subpresheaf of $y_C \times X$. Therefore, $\bar{Z}$ is an element of $\mathcal{P}_{\mathrm{s}}(\mathrm{X})(C)$. We are now going to argue that $\bar{Z}$ is an amalgamation for our compatible family. We have for all $f : D \to C \in R$

$$\begin{array}{llr} & \mathcal{P}_{\mathrm{s}}(\mathrm{X})(f)(\bar{Z}) & \\ = & (y_f \times \mathrm{id}_X)^*(\bar{Z}) & [\text{by definition of } \mathcal{P}_{\mathrm{s}}(\mathrm{X})(f)] \\ = & \overline{(y_f \times \mathrm{id}_X)^*(Z)} & [\text{by (4) in Definition 3.3.13}] \\ = & \bar{Z}_f & [\text{by (3.8)}] \\ = & Z_f. & [\text{since } Z_f \text{ is closed.}] \end{array}$$

So $\bar{Z}$ is indeed an amalgamation as desired. This finishes the proof that $\mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is a sheaf. □

This establishes the action of the powerclass functor $\mathcal{P}_{\mathrm{s}}$ in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ on objects. The explicit description of $\mathcal{P}_{\mathrm{s}}$ on arrows is given as follows. Let $\eta : X \Rightarrow$

Y is an arrow in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ and let for a moment $\mathcal{P} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ be the powerclass functor of $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ that we discussed at the end of Section 3.2. Then $\mathcal{P}_{\mathrm{s}}(\eta)_C : \mathcal{P}_{\mathrm{s}}(X)(C) \to \mathcal{P}_{\mathrm{s}}(Y)(C)$ maps a small and closed subpresheaf $Z \subseteq y_C \times X$ to

$$\mathcal{P}_{\mathrm{s}}(\eta)_C(Z) := \overline{\mathcal{P}(\eta)_C(Z)}.$$

Next, we are going to define the element relation in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$. Let X be a large sheaf. As is the case of presheaves we define $\in_{\mathrm{X}} \subseteq X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X})$ by

$$\in_{\mathrm{X}}(C) = \{(x, A) \in X(C) \times \mathcal{P}_{\mathrm{s}}(\mathrm{X})(C) \mid (\mathrm{id}_C, x) \in A(C)\}.$$

It is easy to check that $\in_{\mathrm{X}}$ defines a subpresheaf. We are going to check that $\in_{\mathrm{X}}$ is a sheaf.

Lemma 3.3.22. *The relation $\in_{\mathrm{X}} \subseteq X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is a subsheaf.*

Proof. Since X is by assumption a sheaf, also $X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is a sheaf. So by Lemma 3.3.17 it suffices to show that $\in_{\mathrm{X}}$ is closed. In order to do so, we will apply Lemma 3.3.16. So let $(x, Z) \in (X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X}))(C)$ such that $S := \{f : D \to C \mid (X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X}))(f)(x, Z) \in \in_{\mathrm{X}}(D)\}$ covers C. We have

$$\begin{aligned}
S = \;& \{f : D \to C \mid (X(f)(x), \mathcal{P}_{\mathrm{s}}(\mathrm{X})(f)(Z)) \in \in_{\mathrm{X}}(D)\} \\
= \;& \{f : D \to C \mid (X(f)(x), (y_f \times \mathrm{id}_X)^*(Z)) \in \in_{\mathrm{X}}(D)\} \quad \text{[by def. of } \mathcal{P}_{\mathrm{s}}(\mathrm{X})(f)] \\
= \;& \{f : D \to C \mid (\mathrm{id}_D, X(f)(x)) \in (y_f \times \mathrm{id}_X)^*(Z)(D))\} \quad \text{[by def. of } \in_{\mathrm{X}}] \\
= \;& \{f : D \to C \mid (f, X(f)(x)) \in (Z)(D))\} \quad \text{[by definition of the pullback)]}
\end{aligned}$$

covers C. Now since Z is a closed subpresheaf of $y_C \times X$, it follows by Lemma 3.3.16 that $(\mathrm{id}_C, x) \in Z(C)$. So by definition of $\in_{\mathrm{X}}$ it follows that $(x, Z) \in \in_{\mathrm{X}}(C)$. □

We will now sketch the proof that $\in_{\mathrm{X}} \subseteq X \times \mathcal{P}_{\mathrm{s}}(\mathrm{X})$ satisfies the universal property. To this end let $R \subseteq X \times Y$ be a small relation in $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$. The unique map $t : Y \Rightarrow \mathcal{P}_{\mathrm{s}}(\mathrm{X})$ is given by the components $t_C : Y(C) \to \mathcal{P}_{\mathrm{s}}(\mathrm{X})(C)$, where for $y \in Y(C)$, $t_C(y)$ is defined by

$$t_C(y)(D) = \{(f, x) \mid f : D \to C, x \in X(D), \text{ such that } (Y(f)(y), x) \in R(D)\}.$$

One can easily check that this defines a subpresheaf of $y_C \times X$. Smallness of $t_C(y)$ follows by R being small. We will check that $t_C(y)$ is closed. Again, we will use the characterization given by Lemma 3.3.16. Assume $f : D \to C$ and $x \in X(D)$ such that

$$\begin{aligned}
& \{g : E \to D \mid (y_C \times X)(f)(g) \in t_C(y)(E)\} \\
= \;& \{g : E \to D \mid (f \circ g, X(g)(x)) \in t_C(y)(E)\} \\
= \;& \{g : E \to D \mid (Y(g)(Y(f)(y)), X(g)(x)) \in R(E)\} \quad \text{[by definition of } t_C(y)(E)]
\end{aligned}$$

covers D. Since R is a sheaf it follows that $(Y(f)(y), x) \in R(D)$. But this shows that $(f, x) \in t_C(y)(D)$ is as desired. This finishes the proof that $t_C(y)$ is closed. One can check that t is the unique map making the desired diagram into a pullback. This shows that (P1) holds. The validity of the powerset axiom follows, as in the case of presheaves, by a precise description of the internal subset relation. Smallness of the natural number object is clear. That finishes the verification of the axioms for the small maps and therefore we have proved Proposition 3.3.20.

3.3.5 The Dense Topology

In this section we will introduce the dense topology, a canonical example for a Grothendieck topology on a small category $\mathcal{C}$. A particular feature of the dense topology is that the resulting large sheaves always form a Boolean pretopos. Let $\mathcal{C}$ be a small category. Let C be in $\mathcal{C}$ and let R be a sieve on C. Let

$$R \text{ be covering } \Leftrightarrow \text{ f.a. } f : D \to C, \text{ there is } g : E \to D \text{ such that } f \circ g \in R.$$

One can easily check that this defines a Grothendieck topology on $\mathcal{C}$. This topology is called the *dense topology* on $\mathcal{C}$ and is often denoted by $\neg\neg$.

Proposition 3.3.23. *Let $\neg\neg$ be the dense topology on $\mathcal{C}$. Then $LSh(\mathcal{C}, \neg\neg)$ is a Boolean pretopos.*

Proof. We need to check that for all $X \in LSh(\mathcal{C}, \neg\neg)$, $\mathrm{Sub}(X)$ is a Boolean algebra, i.e. it remains to show that all $A \in \mathrm{Sub}(X)$ have a Boolean complement. One can calculate that the Heyting complement $\neg A$ of A is explicitly given by

$$\neg A(C) = \{x \mid \text{ f.a. } f : B \to C, X(f)(x) \notin A(B)\}.$$

In order to show that $\neg A$ is a Boolean complement it suffices to show that

$$\neg\neg A \leq A \text{ in } \mathrm{Sub}(X). \tag{3.9}$$

Let $x \in \neg\neg A(C)$. In order to show that $x \in A(C)$ it suffices to show that $R := \{h : D \to C \mid X(f)(x) \in A(D)\}$ covers C in the dense topology since A is a closed subpresheaf of X. So let $f : B \to C$ be arbitrary. Since $x \in \neg\neg A(C)$ it follows that $X(f)(x) \notin \neg A(B)$. But this implies that there is an arrow $g : D \to B$ with $X(g)X(f)(x) = X(f \circ g)(x) \in A(D)$, i.e. $f \circ g \in R$. So we showed that R is covering and therefore (3.9) follows. ☐

3.3.6 Large Sheaves over Poset Categories

For later reference, we will examine the special case where $\mathbb{P}$ is a poset category. In this case most of the basic notions introduced in the preceding sections have a particularly simple form:

Sieves: A sieve on $p \in \mathbb{P}$ corresponds to the downwards closed subsets below p.

The dense topology: A sieve on p, i.e. is a downwards closed subset D below p is covering in the $\neg\neg$-toplogy if and only if D is dense below p, i.e. for every $q \leq p$ there is $d \in D$ with $d \leq q$.

Representable presheaves: For $p \in \mathbb{P}$:

$$\begin{aligned} y_p : & \mathbb{P}^{op} \longrightarrow \text{Classes} \\ & q \mapsto \emptyset \text{ if } p \not\leq q \text{ and } q \mapsto \{*\} \text{ if } p \leq q. \end{aligned}$$

together with the obvious maps.

The $\mathcal{P}_s$-functor on the presheaves over $\mathbb{P}$:

$$\begin{aligned}\mathcal{P}_{\mathrm{s}}(X)(p) =& \{Y \mid Y \subseteq y_p \times X \text{ small subpresheaf } \}\\ \cong& \{Y \mid Y \subseteq X \text{ small subpresheaf such that } Y(q) = \emptyset \text{ if } q \not\leq p\}\end{aligned}$$

Also, for $p' \leq p$, the map

$$\mathcal{P}_{\mathrm{s}}(X)(p' \leq p) : \mathcal{P}_{\mathrm{s}}(X)(p) \longrightarrow \mathcal{P}_{\mathrm{s}}(X)(q)$$

is given by the obvious inclusion. For this reason we will often abuse notation by writing for example $x \in (y_p \times X)(q)$ for some $x \in X(q)$.

Given these descriptions we can easily infer the following corollaries that we establish for later reference. By Remark 3.2.4, we know that the large presheaves over $\mathbb{P}$ are generated by the representable presheaves. One can infer that the large sheaves are therefore generated by the collection $\mathbf{a}(y_p)$, $p \in \mathbb{P}$, see for example the discussion in [20, p. 139]. The above description of the representable presheaves reveals that these are all subobjects of 1, the terminal presheaf. Since the sheafification functor $\mathbf{a}$ preserves the terminal object and monos we can infer that also the large sheaves are generated by subobjects of 1. This gives:

Corollary 3.3.24. *$LSh(\mathbb{P}, \neg\neg)$ is generated by subobjects of 1.*

By Proposition 3.3.23 from above it follows that $LSh(\mathbb{P}, \neg\neg)$ is Boolean. Since as in the case of the presheaves one can sow that all even class-sized limits and colimts exist, we can infer that the subobject lattice of a large sheaf is complete. Therefore, we also established:

Corollary 3.3.25. *$LSh(\mathbb{P}, \neg\neg)$ is Boolean pretopos. Moreover, for every X, $\mathrm{Sub}(X)$ is a complete Boolean algebra.*

Chapter 4

Algebraic Universes of Set Theory

Algebraic universes of set theory arise as initial algebras for the powerclass functor on a category of classes. In this chapter we will examine these initial algebras. As first shown in [18], we will prove that they satisfy an intuitionistic version of set theory. Moreover, we will investigate the correspondence between a category of classes and the universe that it contains.

In the first section, we will explain how to interpret the language of set theory in an algebraic universe and introduce the axioms of the set theory **IZF**. In the second section we will show that our axiomatization of a category of classes provides universes for the set theory **IZF**, and for **ZF** in the case the underlying pretopos is Boolean. Moreover, we will examine the validity of the axiom of choice in our framework. In the third section we will consider more elementary set theoretical constructions in the universe. In particular, we will see how set theoretical structure provided by the Heyting pretopos can be "squeezed" into the universe that it contains. The established results provide us the necessary tools for our considerations in the next chapter. In the last section we will investigate the existence of initial algebras for the $\mathcal{P}_{\mathrm{s}}$-functor.

4.1 Interpreting the Language of Set Theory

In this section, we will see how to interpret the language of set theory in a category of classes that contains an initial algebra for the $\mathcal{P}_{\mathrm{s}}$-functor.

The language of set theory as a sorted language consists of:

- Variables that all have the same sort V.
- A binary relation symbol $\in : V \times V$.

Let $\mathcal{E}$ be a category of classes and assume that V is an initial algebra for the powerclass functor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$. The unique sort of the language of set theory will be interpreted by V. We will now show how to interpret the element relation. Since V is initial, by Lambek's Lemma we get isomorphisms:

$$\begin{tikzcd} \mathcal{P}_{\mathrm{s}}(V) \arrow[r, bend left, "\mathrm{Int}"] & V \arrow[l, bend left, "\mathrm{Ext}"] \end{tikzcd}$$

Recall that by axiom (P1) there is a small relation $\in_V$ on $V \times \mathcal{P}_{\mathrm{s}}(V)$. By taking the pullback

$$\begin{tikzcd} \varepsilon \arrow[r] \arrow[d, tail] & \in_V \arrow[d, tail] \\ V \times V \arrow[r, "\mathrm{id}_V \times \mathrm{Ext}"'] & V \times \mathcal{P}_{\mathrm{s}}(V) \end{tikzcd}$$

we obtain a small relation ε on $V \times V$. The relation ε provides a canonical interpretation of the language of set theory in $\mathcal{E}$. If $\varphi(x_1, \ldots, x_n)$ is a formula in the language of set theory, such that all free variables of φ are among $\bar{x} = x_1, \ldots, x_n$ then we will write

$$(V, \varepsilon) \models \varphi \text{ iff } [\![\bar{x} \mid \varphi]\!] \text{ is the maximal in } \mathrm{Sub}(\prod_{1 \leq l \leq n} V).$$

Remark 4.1.1. Note that alternatively, we could have defined the interpretation of $\in$ as the subobject $\exists_{\mathrm{id}_V \times \mathrm{Int}}(\in_V)$ of $V \times V$, i.e. as the composite $\in_V \to V \times \mathcal{P}_{\mathrm{s}}(V) \to V \times V$. But since Ext and Int are inverse to each other, these two canonical choices deliver the same relation.

Note that since Ext and Int are inverse to each other, also

$$\begin{tikzcd} \in_V \arrow[r] \arrow[d, tail] & \varepsilon \arrow[d, tail] \\ V \times \mathcal{P}_{\mathrm{s}}(V) \arrow[r, "\mathrm{id}_V \times \mathrm{Int}"] & V \times V \end{tikzcd} \tag{4.1}$$

is a pullback. It is now easy to see that the representability axiom (P1) for small relations on $\in_V$ translates to one on ε.

($P1'$) Let $r : R \to V \times X$ be a small relation. Then there exists a unique arrow $t : X \to V$ that fits into a pullback:

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] & \varepsilon \arrow[d, tail] \\ V \times X \arrow[r, "\mathrm{id}_V \times t"'] & V \times V \end{tikzcd}$$

Moreover, if t' classifies R is terms of $\in_V$, then $t = \mathrm{Int} \circ t'$.

Clearly, as in Section 2.3, a small relation $r : R \to V \times X$ can be characterized internally as

$$\mathcal{E} \models_{x:X, v:V} Rxv \leftrightarrow x \varepsilon t(y) \leftrightarrow x \in_V t'(y).$$

In particular, if $m : A \to V$ is a small subobject then according to (P1$'$) it can be classified by a global section $a : 1 \to V$, i.e. we have a pullback

$$\begin{tikzcd} A \arrow[r] \arrow[d, tail, "m"'] & \varepsilon \arrow[d, tail] \\ V \times 1 \arrow[r, "\mathrm{id}_V \times a"'] & V \times V \end{tikzcd}$$

Note that in this case A and $[\![x : V \mid x\varepsilon a]\!]$ are the same subobjects of V, i.e. we have:

$$\begin{tikzcd} A \arrow[rr, dash, "\cong"] \arrow[dr, tail, "m"'] & & {[\![x : V \mid x\varepsilon a]\!]} \arrow[dl, tail] \\ & V & \end{tikzcd} .$$

This correspondence tells us that small subobjects of the universe correspond to elements of the universe.

Remark 4.1.2. Analogously to usual practice in model theory, we will use the following conventions: We will allow to use function symbols in formulas in the language of set theory for every arrow $f : V^n \to V$ of $\mathcal{E}$. So in particular, if $a : 1 \to V$ is a global sections we regard $x\varepsilon a$ as a valid formula in the language of set theory.

Recall that in Section 2.2 we defined the internal subsetrelation $\subseteq_V$ on $\mathcal{P}_{\mathrm{s}}(V) \times \mathcal{P}_{\mathrm{s}}(V)$. Now, we can also form a the subsetrelation on V via

$$\subseteq := [\![x, y : V \mid \forall z(z\varepsilon x \to z\varepsilon y)]\!] \longrightarrow V \times V.$$

The difference between the two relations is that for $\subseteq_V$, we used the "typed version" $\in_V$ of the element relation whereas in the latter we used the "untyped" ε. By the pullback (4.1), the preservation of Boolean operations under pullback and since the quantification satisfy the Beck-Chevalley condition, we obtain that $\subseteq_V$ is a pullback of $\subseteq$.

Lemma 4.1.3. *The following is a pullback*

$$\begin{tikzcd} \subseteq_V \arrow[r] \arrow[d, tail] & \subseteq \arrow[d, tail] \\ \mathcal{P}_{\mathrm{s}}(V) \times \mathcal{P}_{\mathrm{s}}(V) \arrow[r, "\mathrm{Int} \times \mathrm{Int}"'] & V \times V \end{tikzcd} .$$

Remark 4.1.4. Of course, the above correspondence is true more generally. Assume that ϕ is a formula in the internal language of $\mathcal{E}$ containing variables of types ${\mathcal{P}_{\mathrm{s}}}^n(V)$ for some n, using equality and the element relations $\in_{{\mathcal{P}_{\mathrm{s}}}^n(V)}$. Then by basically ignoring the types of the variables, we can interpret the formula ϕ using the relation ε. We would call this the untyped interpretation of ϕ. By iteratively applying the isomorphism Int with inverse Ext we obtain, similar as above, a pullback diagram between the typed and the untyped interpretation of ϕ.

4.1.1 The Constructive Set Theory IZF

The development of constructive set theory was initiated by John Myhill in his seminal paper [22]. Myhill's intention was to give Bishop's program on constructive mathematics a set theoretical foundation. Since then many axiomatizations of constructive set theories were introduced and studied. See for example Peter Aczel's and Michael Rathjen's book [1] for a good introduction to the topic. The main features that constructive set theories opposed to classical Zermelo Fraenkel set theory have are:

- The underlying logic follows intuitionistic rather than classical rules.

Therefore,

- the axiom of choice must be dropped, as (in the presence of the other axioms) it implies the law of excluded the middle by a famous result of Diaconescu[1].
- In the presence of the other axioms, the foundation axiom implies the law of excluded in the middle. Therefore the axiom of foundation is replaced by the classically equivalent axiom scheme of set induction.

Accordingly, we define the theory **IZF** that we will refer to as intuitionistic set theory as follows.

Definition 4.1.5. The axioms of intuitionistic set theory **IZF** are:

(**Extensionality**): $\forall x \forall y((\forall z(z \in x \leftrightarrow z \in y)) \leftrightarrow x = y)$

(**Pairing**): $\forall x \forall y \exists z \forall v(v \in z \leftrightarrow ((v = x) \vee (v = y)))$

(**Union**): $\forall x \exists z \forall v(v \in z \leftrightarrow \exists w(w \in x \wedge v \in w))$

(**Powerset**): $\forall x \exists z \forall v(v \in z \leftrightarrow \forall w(w \in v \rightarrow w \in x))$

(**Infinity**) : $\exists x(\emptyset \in x \wedge (y \in x \rightarrow y \cup \{y\} \in x))$

(**Separationscheme**): For every formula φ, where a does not occur freely: $\forall a \exists w \forall x(x \in w \leftrightarrow (x \in a \wedge \varphi))$

(**Replacementscheme**): For every formula φ, where w does not occur freely: $(\forall x\,(x \in a \rightarrow \exists! y \varphi(x, y))) \rightarrow \exists w \forall y(y \in w \leftrightarrow \exists x(x \in a \wedge \varphi))$

(**Induction**): $\forall x(\forall y(y \in x \rightarrow \varphi) \rightarrow \varphi[x/y]) \rightarrow \forall x \varphi$

4.2 Algebraic Models for (Intuitionistic) Set Theories.

We will now prove that algebraic universes satisfy the axioms of **IZF** introduced in the previous section. As already mentioned, the validity of the particular set theoretic axioms relies on the axioms for small maps that we imposed on a category of classes. We will point out the precise correspondences in the proof

[1]See for example [1, p. 104] for a proof of this fact.

of Theorem 4.2.2. Whereas the axioms for the small maps determine which of the set theoretical axioms are validated, the internal logic of the underlying pretopos determines the underlying logic of the set theory. In Proposition 4.2.3, we will show that a Boolean pretoposes carrying the structure of a category of classes gives rise to a universe of classical set theory, i.e. satisfying the usual **ZF** axioms and is sound with respect to classical inference rules.

The following Lemma will be useful in the proof of Theorem 4.2.2. It expresses that if a formula defines a small relation on V, then the universe believes that there is a set containing exactly those elements validating the formula.

Lemma 4.2.1. *Let φ be a formula in the language of set theory such that the free variables of φ are among y and $\bar{x}$. Assume the interpretation*

$$[\![y\bar{x}.\varphi]\!] \xrightarrow{m} V \times V^{\bar{x}}$$

defines a relation that is small in V, where $V^{\bar{x}} := \prod_{x \in \bar{x}} V$. Then

$$(V, \varepsilon) \models \exists w(y \varepsilon w \leftrightarrow \varphi).$$

Proof. Since m as a relation is small in V, by (P1$'$) there is a unique map $f : V^{\bar{x}} \to V$ fitting in the pullback

$$\begin{tikzcd}
{[\![y\bar{x}.\varphi]\!]} \arrow[r] \arrow[d, tail, "m"'] & \varepsilon \arrow[d, tail] \\
V \times V^{\bar{x}} \arrow[r, "\mathrm{id}_V \times f"] & V \times V
\end{tikzcd}.$$

It is clear that

$$\mathcal{E} \models y \varepsilon f(\bar{x}) \leftrightarrow \varphi(y, \bar{x}).$$

The rest of the proof follows by Lemma 2.1.17. $\square$

Theorem 4.2.2. *Let $\mathcal{E}$ be a category of classes and let V be an initial algebra for the $\mathcal{P}_{\mathrm{s}}$ functor on $\mathcal{E}$. Then,*

$$(V, \varepsilon) \models \textbf{IZF}$$

Proof. We check that (V, ε) validates the axiom of **IZF** listed in Definition 4.1.5.

(**Extensionality**): The proof is the same as that of typed extensionality in Lemma 2.3.3. One only needs to replace the use of axiom (P1) by axiom (P1$'$).

Checking the validity of the axioms of pairing, union, and powerset is very similar. Each of these axioms requires the existence of a set defined by a certain formula. So provided that the formula defines a small relation, we obtain a set as in Lemma 4.2.1. The smallness of the relations in question relies on a corresponding axiom for the small maps.

(**Pairing**): For the case of the pairing axiom, we need to show that $[\![z = x \vee z = y]\!] \to V \times V \times V$ is a small relation. Then as in Lemma 4.2.1, we obtain an arrow

$$\mathbf{pair} : V \times V \longrightarrow V \text{ such that}$$

$[\![z \in \mathbf{pair}(x,y) \leftrightarrow z = x \vee z = y]\!] = V \times V \times V$ in $\mathrm{Sub}(V \times V \times V)$ which implies the validity of the pairing axiom. Let V_x, V_y and V_z denote the copies of V intended to interpret the variables x, y and z, respectively.

First note that

$$[\![zxy.z = x]\!] \cong V \times V \xrightarrow{\langle \pi_1, \pi_1, \pi_2 \rangle} V_z \times (V_x \times V_y) \quad \text{and}$$

$$[\![zxy.z = y]\!] \cong V \times V \xrightarrow{\langle \pi_1, \pi_1, \pi_2 \rangle} V_z \times (V_x \times V_y)$$

define small relations in V_z.

By axiom (S5) , it follows that the composite

$$[\![zxy.z = x]\!] + [\![zxy.z = y]\!] \longrightarrow V_z \times (V_x \times V_y) \longrightarrow V_x \times V_y$$

is a small map. Now the interpretation $[\![zxy.z = x \vee z = y]\!]$ is the image in

$$\begin{tikzcd}
{[\![zxy.z = x]\!] + [\![zxy.z = y]\!]} \arrow[r] \arrow[dr, two heads] & V_z \times (V_x \times V_y) \arrow[r] & V_x \times V_y \\
& {[\![zxy.z = x \vee z = y]\!]} \arrow[u, tail, "m"'] &
\end{tikzcd}$$

.

And therefore, $m : [\![zxy.z = x \vee z = y]\!] \to V_z \times (V_x \times V_y)$ is a small relation by (S3) as desired.

(**Union**)**:** We need to show that $[\![xw.\exists y(y\varepsilon x \wedge w\varepsilon y)]\!] \to V \times V$ defines a small relation. Then as before, the validity of the union axiom follows by Lemma 4.2.1. As in the proof of Lemma 2,

$$[\![xyw.y\varepsilon x \wedge w\varepsilon y]\!] \rightarrowtail (V_y \times V_w) \times V_x \tag{4.2}$$

defines a relation that is small in $V_y \times V_w$. Now $[\![xw.\exists y(y\varepsilon x \wedge w\varepsilon y)]\!]$ is the image in

$$\begin{tikzcd}
{[\![xyw, y \in x \wedge w\varepsilon y]\!]} \arrow[r] \arrow[dr, two heads] & V_y \times V_w \times V_x \arrow[r] & V_x \times V_w \\
& {[\![xw.\exists y(y\varepsilon x \wedge w\varepsilon y)]\!]} \arrow[ur, tail] &
\end{tikzcd}$$

Using that $[\![xyw.y \in x \wedge w \in y]\!] \to V_x$ is small by 4.2 and axiom (S3) it follows that

$$[\![xw.\exists y(y \in x \wedge w \in y)]\!] \rightarrowtail V_w \times V_x$$

is a small relation as desired.

(**Powerset**)**:** We need to construct a map

$\mathbf{pow} : V \to V$ such that $[\![v \in \mathbf{pow}(x) \leftrightarrow \forall w(w \in v \to w \in x))]\!]$ in $\mathrm{Sub}(V \times V)$.

Axiom (P2) says that

$$\subseteq_V := [\![\forall w : V(w \in v \to w \in x)]\!] \rightarrowtail \mathcal{P}_{\mathrm{s}}(V) \times \mathcal{P}_{\mathrm{s}}(V)$$

is a small relation. Following Remark 4.1.4 at the beginning of the current chapter, and using how ε is defined in terms of $\in_V$, the internal subsetrelation on V can be obtained as the pullback

$$\begin{tikzcd}
[\![\forall w(w \in v \to w \in x)]\!] \arrow[r] \arrow[d] & \subseteq_V \arrow[d] \\
V \times V \arrow[r, "\mathrm{Ext} \times \mathrm{Ext}"] & \mathcal{P}_{\mathrm{s}}(V) \times \mathcal{P}_{\mathrm{s}}(V)
\end{tikzcd}.$$

Since Ext is an isomorphism, it follows by (P2) that $[\![\forall w(w \in v \to w \in x)]\!]$ defines a small relation on V. The map **pow** can be defines as the classifying map of this relation.

(Separation): The validity of the axiom of separation is an immediate consequence of all monos being small.

Assume $\varphi(y, \bar{x})$ is a formula where a does not occur freely and such that all free variables of φ are among the ones in $y, \bar{x}$. Consider the diagram

$$\begin{tikzcd}
[\![ya\bar{x}.y \in a \wedge \varphi(y, \bar{x})]\!] \arrow[r, tail, "n"] \arrow[d, tail] & [\![ya\bar{x}.\, y\varepsilon a]\!] \arrow[d, tail, "k"] \\
[\![ya\bar{x}.\, \varphi(y, \bar{x})]\!] \arrow[r, tail] & V_y \times V_a \times V^{\bar{x}} \arrow[d, "\pi"] \\
 & V_a \times V^{\bar{x}}
\end{tikzcd},$$

where the upper square is a pullback so that $k \circ n : [\![ya\bar{x}.y \in a \wedge \varphi(y, \bar{x})]\!] \to V_y \times V_a \times V^{\bar{x}}$ provides the interpretation of $y \in a \wedge \varphi(y, \bar{x})$. Since ε is a small relation one can easily infer that the composite on the right had side is a small map. As a pullback of a mono, the map n is a mono and therefore small. It follows that the composite $\pi k n$ is a small map. And so $[\![ya\bar{x}.y \in a \wedge \varphi(y, \bar{x})]\!] \to V_y \times V_a \times V^{\bar{x}}$ is small in V_y as desired.

(**Infinity**) : The validity of the axiom of infinity relies on the smallness of the natural number object. We will show that the recursive property of the natural number object implies that it is a subobject of V. We define a successor operation $\mathbf{succ} : V \to V$, intended to map a set x to $x \cup \{x\}$ as the composite

$$V \xrightarrow{(\mathrm{id}_V, s)} V \times V \xrightarrow{\mathbf{pair}} V \xrightarrow{\mathbf{union}} V,$$

where **pair** and **union** are defined as previously and $s : V \to V$ is the singleton map, the classifying map of the diagonal on V. The classifying map $\emptyset : 1 \to V$ of the unique map $0 \to V \times 1$ from the initial object, defines the empty set in V. We show that there is a map $\omega : 1 \to V$ such that

$$(V, \varepsilon) \models \emptyset \varepsilon \omega \wedge \forall x (x \varepsilon \omega \to \mathbf{succ}(x) \varepsilon \omega).$$

By the recursive property of the natural number object in $\mathcal{E}$, there is a unique map $g : \mathbf{N} \to V$ that fits into a commutative diagram:

$$\begin{tikzcd}
1 \arrow[r, "\mathbf{o}"] \arrow[d, "\cong"] & \mathbf{N} \arrow[r, "\mathbf{s}"] \arrow[d, "g"'] & \mathbf{N} \arrow[d, "g"] \\
1 \arrow[r, "\emptyset"] & V \arrow[r, "\mathbf{succ}"] & V
\end{tikzcd}.$$

Arguing internally, one can show that g is a mono, so that $g : \mathbf{N} \to V$ defines a small subobject of V. Let $\omega : 1 \to V$ be its classifying map. Then clearly,

$$\mathcal{E} \models_{x:V} x \varepsilon \omega \leftrightarrow \exists n : \mathbf{N}\, g(n) = x. \tag{4.3}$$

Using (4.3) and the commutativity of the diagram above, we can easily prove that ω is as required. So in particular:

$$\mathcal{E} \models g(\mathbf{o}) = \emptyset \to \exists n : \mathbf{N}\, g(n) = \emptyset \to \emptyset \varepsilon \omega.$$

and

$$\begin{aligned}
\mathcal{E} \models x \varepsilon \omega &\to \exists n : \mathbf{N}\, g(n) = x \\
&\to \exists n : \mathbf{N}\, g(\mathbf{s}(n)) = \mathbf{succ}(x) \\
&\to \exists m : \mathbf{N}\, g(m) = \mathbf{succ}(x) \\
&\to \mathbf{succ}(x) \varepsilon \omega.
\end{aligned}$$

(**Induction**) : For readability reasons, we first discuss the case where $\varphi(x)$ is a formula in the language of set theory with only one free variable. First assume that

$$T := [\![\forall y \forall x (x \varepsilon y \to \varphi(x)) \to \varphi(y)]\!] = 1 \text{ in } \mathrm{Sub}(1).$$

We will later argue how the general case follows from that. By the above, in particular $[\![\forall x (x \varepsilon y \to \varphi(x)) \to \varphi(y)]\!] = V$ in $\mathrm{Sub}(V)$. We want to show that $[\![\varphi(x)]\!] = V$. To this end, we would like to equip $m : A := [\![\varphi]\!] \to V$ with a $\mathcal{P}_{\mathrm{s}}$-algebra structure. Then the claim follows by the inductive property of the initial algebra discussed in Section 2.3.3. Define

$$R := [\![w : \mathcal{P}_{\mathrm{s}}(A) y : A \mid m(y) = \mathrm{Int}(\mathcal{P}_{\mathrm{s}}(m)(w))]\!] \to \mathcal{P}_{\mathrm{s}}(A) \times A.$$

We claim that R defines a functional relation on $\mathcal{P}_{\mathrm{s}}(A) \times A$. The functionality of the relation is clear by the definition of R. It remains to show that R is total. Observe that to this end, we need to show that $[\![w : \mathcal{P}_{\mathrm{s}}(A) \mid \varphi(\mathrm{Int}(\mathcal{P}_{\mathrm{s}}(m)(w)))]\!] \cong \mathcal{P}_{\mathrm{s}}(A)$ in $\mathrm{Sub}(A)$. Note that by our assumption (4.2), it suffices to show that $\mathcal{E} \models x \varepsilon \,\mathrm{Int}(\mathcal{P}_{\mathrm{s}}(m)(w)) \to \varphi(x)$. But we have

$$\begin{aligned} \mathcal{E} \models x\varepsilon \,\mathrm{Int}\, \mathcal{P}_{\mathrm{s}}(m)(w) &\to x \in \mathcal{P}_{\mathrm{s}}(m)(w) && [\text{By definition of } \varepsilon] \\ &\to \exists z : A\,(m(z) = x \wedge z \in_A w) && [\text{By (2.3), Lemma 2.3.4.}] \\ &\to \varphi(x) && , \end{aligned}$$

where the last implication follows since $[\![\varphi(x)]\!] = [\![\exists z : A m(z) = x]\!]$ in $\mathrm{Sub}(V)$. This finishes the proof that R is a functional relation. Now by Lemma 2.1.16, of Section 2.1.2, R corresponds to an arrow $\alpha : \mathcal{P}_{\mathrm{s}}(A) \to A$ with graph R. It is immediate to see that f equips A with an algebra structure such that $m : A \to V$ becomes an morphism of $\mathcal{P}_{\mathrm{s}}$-algebras. As explained above, we can conclude that $[\![\varphi(y)]\!] = V$. Next, we will explain how we can get rid of the assumption (4.2). In the slice category T, the interpretation of the formula $\forall y \forall x (x \varepsilon y \to \varphi(x)) \to \varphi(y)$ is the maximal subobject of the terminal object in $\mathcal{E}\,/T$. Note that by our considerations in Section 2.3.3, we can infer that $\langle \mathrm{Int}, \mathrm{id}_T \rangle \mathcal{P}_{\mathrm{s}}(V) \times T \to V \times T$, the pullback of the algebra $\mathcal{P}_{\mathrm{s}}(V) \to V$ along $t : T \to 1$, gives rise to a ${\mathcal{P}_{\mathrm{s}}}^T$-algebra satisfying the inductive property. Therefore, we can carry out the proof above in the category $\mathcal{E}\,/T$, to obtain that $[\![\forall y(\varphi(y)) \cong T$. This finishes the proof that (V, ε) validates the induction axiom for formulas φ with one free variable. Now if φ contains more free variables, we can carry out the same proof by passing to the appropriate slice category.

(**Replacement**) : This proof is taken from [30]. First assume that $a : 1 \to V$ is a global section such that $[\![\forall x(x\varepsilon a \to \exists ! y \varphi(x,y))]\!] = 1$ in $\mathrm{Sub}(1)$, i.e. $(V,\varepsilon) \models \forall x(x\varepsilon a \to \exists ! y\varphi(x,y))$. Let $A \to V$ be the small subobject of V that is characterized by a. Define a relation

$$R := [\![x : V y : V \mid \varphi(x,y) \wedge x\varepsilon a]\!] \to A \times V.$$

By the assumption it is clear that R is functional. So by Lemma 2.1.16, R corresponds to a function $f : A \to V$ such that R is the graph of f. The image of f, is given by the mono n in

$$\begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, "e"'] & & V \arrow[r, "!_V"'] & 1 \\ & \mathrm{Im}(f) \arrow[ur, tail, "n"'] & & \end{tikzcd}.$$

Since A is a small, the composite $!_V \circ f$ is a small map. Since e is epi, by (S3) it follows that also $!_V \circ n$ is a small subobject of V. Note that $\mathrm{Im}(f) = [\![\exists x : V x\varepsilon a \wedge \varphi(x,y)]\!]$. Since $\mathrm{Im}(f)$ is small, using Lemma 4.2.1, it follows that $(V,\varepsilon) \models \exists w(\forall x(x\varepsilon w \leftrightarrow (x\varepsilon a \wedge \varphi(x,y)))$. The general case follows by passing to the appropriate slice category and using the stability under slicing results that we discussed in Section 2.3.2.

□

Now, if $\mathcal{E}$ is a category of classes such that the underlying pretopos is Boolean, then the internal logic of $\mathcal{E}$ follows classical rules. In Section 4.1.1 we already mentioned that the axiom of set induction is classically equivalent to the axiom of foundation. Therefore, in this case our algebraic universe satisfies the usual **ZF** axioms.

Corollary 4.2.3. *If the underlying pretopos $\mathcal{E}$ of the category of classes is Boolean then*

$$(V, \varepsilon) \models \mathbf{ZF}\,.$$

4.2.1 About The Axiom of Choice

We would like to say a few words about the axiom of choice in our context. In topos theory, the axiom of choice is usually formulated in the form that every epi $p : X \to I$ has a section, i.e. there is a morphism $s : I \to X$ such that $p \circ s = \mathrm{id}_I$. However, the validity of the axiom of choice in the internal logic of a topos depends on a weaker principle, called (IAC) stating that for every object E of the topos, the functor $(-)^E : \mathcal{E} \to \mathcal{E}$ preserves epimorphisms. See for example the discussion in [20, p. 275]. We say that a category of classes $\mathcal{E}$ satisfies the axiom of choice, if for every small object E an epi $p : E \to I$ with domain E has a section. The internal version translates to the statement that for every small object E, the functor $(-)^E : \mathcal{E} \to \mathcal{E}$ preserves epimorphisms between small objects. Note that the latter makes sense in a category of classes as small objects are exponentiable as we discussed in Section 2.3.5. If $\mathcal{E}$ satisfies (IAC) and V is a universe of $\mathcal{E}$ one can prove along the lines of [20, pp. 312-315] that the axiom of choice holds in the universe.

Lemma 4.2.4. *If a category of classes $\mathcal{E}$ satisfies (IAC), then $(V, \varepsilon) \models AC$.*

Assuming the axiom of choice in our metatheory, one can prove that a Boolean topos that is generated by subobjects of 1 and whose subobject lattices are complete Boolean algebras satisfies the axiom of choice, see Proposition 8 in [20, p. 276]. The idea of the proof is as follows. Assuming that $p : E \to I$ is epi, we can consider the partial order of partial sections of p, i.e. of pairs (W, s) with $m : W \to I$ in $\mathrm{Sub}(I)$ and $s : W \to E$ such that $e \circ s = m$. By completeness of the subobject lattices, we can apply Zorn's Lemma to obtain a maximal such subobject. Booleanness and the assumption on the generators helps to show that this subobject is I itself.

The criterion can be translated into our framework. However, we need to be a bit careful about sizes in our metatheory. In addition to Booleanness of a category of classes and the assumption on the generators, we need to require that for every small object X the collection of subobjects of X forms a set in our metatheory; moreover, we need to assume that also the collection of maps between two small object forms a set. In this case the above proof translates immediately. Assume that $\mathcal{E}$ is such a category of classes. Let E be small and assume that $p : E \to I$ is epi. Then by (S3) also I is a small of $\mathcal{E}$. By our assumption on the sizes, the order of partial sections for p forms a set. Therefore we can apply Zorn's Lemma as in the above argument.

Proposition 4.2.5. *Let $\mathcal{E}$ be category of classes such that the underlying pretopos in Boolean. Assume that it is generated by subobjects of* 1 *and assume that for every X is $\mathcal{E}$ the Boolean algebra* $\mathrm{Sub}(X)$ *is complete. Moreover, assume that for every small object X the collection of subobjects of X is a set and that for two small objects X, Y the collection of maps between X and Y is a set. Then $\mathcal{E}$ satisfies the axiom of choice.*

Recall our discussion on a category of large sheaves over a poset in Section 3.3.6. Assume that $LSh(\mathbb{P}, \neg\neg)$ is the category of large sheaves over some poset category $\mathbb{P}$. In Corollary 3.3.24 we showed that it is generated by subobjects of 1. Moreover, in Corollary 3.3.25, we argued that $LSh(\mathbb{P}, \neg\neg)$ is a Boolean pretopos and for every X the subobject lattice $\mathrm{Sub}(X)$ is a complete Boolean algebra. So we can apply the above Proposition 4.2.5 and obtain:

Corollary 4.2.6. $LSh(\mathbb{P}, \neg\neg)$ *satisfies the axiom of choice.*

4.3 More Set theoretic Structure in the Algebraic Universe

In the previous section we saw that the algebraic universes validate the **IZF** axioms. This in particular implies that we can perform elementary set theoretic operations such as products, powersets and function spaces in our universe. The goal of this section is to show that in fact we can perform the same operations outside the universe by using the structure given by the Heyting pretopos and then "squeeze" the obtained object into the universe. We will explain this a bit more carefully. Recall that at the end of Section 4.1, we argued that every small subobject of the universe corresponds to an element of the universe. Assume now that $A \to V$ and $B \to V$ are such small subobjects, corresponding to elements $a : 1 \to V$ and $b : 1 \to V$. Now, we will see that also $A \times B$, $\mathcal{P}_{\mathrm{s}}(A)$ and A^B the product, powerset and exponential of two small subobjects give rise to small subobjects of V in such a way, that V believes that their classifying map define product, powerset and function space between a and b, respectively. One can interpret this correspondence as telling us that statements in the internal logic of the "outer" Heyting pretopos can be translated into statements about the universe. Very similar ideas can also be found in [30], [28], and [6].

In set theory, we code products, functions etc. by ordered pairs. The next Lemma shows that the same such coding can be performed in the algebraic universe.

Lemma 4.3.1. *There is a mono* $V \times V \xrightarrow{\langle \text{-}, \text{-} \rangle} V$ *that corresponds to mapping elements of V to their ordered pair.*

Proof. We simply mimic the Kuratowski definition of the ordered pair in our algebraic universe. Therefore, we define $\langle \text{-}, \text{-} \rangle$ as the composite

$$V \times V \xrightarrow{\Delta_V \times \mathrm{id}_V} V \times V \times V \xrightarrow{s \times \mathbf{pair}} V \times V \xrightarrow{\mathbf{pair}} V \;,$$

where s is the singleton operation and **pair** is the pairing function as in the proof of Theorem 4.2.2. We can infer that

$$(V, \varepsilon) \models z\varepsilon\langle x, y\rangle \leftrightarrow z = s(x) \vee z = \mathbf{pair}(x, y).$$

Using the internal logic, we can prove that

$$(V, \varepsilon) \models \langle x, y\rangle = \langle x', y'\rangle \to (x = x' \wedge y = y').$$

That proves the claim. $\square$

The following Lemma helps us to translate statements of the (typed) internal language of $\mathcal{E}$ into the language of set theory.

Lemma 4.3.2. *Let $\varphi(\bar{x}, y)$ be a formula in the internal language of V such that all the free variables of φ are of type V. Let $m : A \to V$ be a small subobject classified by $a : 1 \to V$, we have:*

$$\llbracket \bar{x} \mid \exists z : A\varphi[m(z)/y] \rrbracket = \llbracket \bar{x} \mid \exists v : V(v\varepsilon a \wedge \varphi)\rrbracket \tag{4.4}$$

and

$$\llbracket \bar{x} \mid \forall z : A\varphi[m(z)/y] \rrbracket = \llbracket \bar{x} \mid \forall v : V(v\varepsilon a \to \varphi)\rrbracket \tag{4.5}$$

in $\mathrm{Sub}(V_{\bar{x}})$*, where $V_{\bar{x}} := \prod_{x\in\bar{x}} V$. Also a more general version of the statement holds. Let $\varphi(\bar{x}, \bar{y})$ be a formula in the internal language with all free variables of type V and let k be the length of the string $\bar{y}$. Let $(m_i : B_i \to V)$, $1 \leq i \leq k$ be small subobjects classified by maps $b_i : 1 \to V$, $1 \leq i \leq k$, then*

$$\llbracket \bar{x} \mid \exists z_1 : B_1 \dots \exists z_k : B_k\, \varphi\left[\overline{m(z_i)}/\bar{y}\right] \rrbracket = \llbracket \bar{x} \mid \bigwedge_i \exists v_i : V v_i\varepsilon b_i \wedge \varphi\rrbracket$$

and

$$\llbracket \bar{x} \mid \forall z_1 : B_1 \dots \forall z_k : B_k\varphi\left[\overline{m(z_i)}/\bar{y}\right] \rrbracket = \llbracket \bar{x} \mid \forall v_1 : V \dots v_k : V \bigwedge_i v_i\varepsilon b_i \to \varphi\rrbracket$$

in $\mathrm{Sub}(V_{\bar{x}})$.

Proof. We show (4.4). Consider the following diagram, where the upper square is a pullback and $n \circ e$ is the epi-mono factorization of $p' \circ l'$ so that $n : T \to V_{\bar{x}}$ is the interpretation of $\exists z : A\varphi[m(z)/y]$.

$$\begin{tikzcd}
 & \llbracket \varphi[m(z)/y] \rrbracket \arrow[r] \arrow[dl, "e"'] \arrow[d, tail, "l'"] & \llbracket \varphi(\bar{x}, v)\rrbracket \arrow[d, tail, "l"] \\
T \arrow[dr, tail, "n"'] & V_{\bar{x}} \times A \arrow[r, tail, "\mathrm{id}_{V_{\bar{x}}} \times m"] \arrow[d, "p'"] & V_{\bar{x}} \times V \arrow[d, "p"] \\
 & V_{\bar{x}} \arrow[r, dash, "\mathrm{id}_{V_{\bar{x}}}"] & V_{\bar{x}}
\end{tikzcd}$$

By our observation at the beginning of this chapter, $m : A \to V$ is the interpretation of $v\varepsilon a$. Therefore, $(\mathrm{id}_{V_{\bar{x}}} \times m) \circ l' : [\![\varphi[m(z)/y]\,]\!] \to V_{\bar{x}} \times V$ is the interpretation of $v\varepsilon a \wedge \varphi$. Now $n \circ e$ is also the epi-mono factorization of $p \circ (\mathrm{id}_{V_{\bar{x}}} \times m) \circ l'$. But this shows that $n : T \to V_{\bar{x}}$ is the interpretation of $\exists v : V(v\varepsilon a \wedge \varphi)$. So we established (4.4). For (4.5), note that

$$[\![\bar{x} \mid \forall z : A\varphi[m(z)/y]\,]\!] = \forall_{p'}(\mathrm{id}_{V_{\bar{x}}} \times m)^{-1}(l : [\![\varphi[m(z)/y]\,]\!] \to V_{\bar{x}} \times V)$$

and

$$[\![\bar{x} \mid \forall v : V(v\varepsilon a \to \varphi]\!] = \forall_p \forall_{\mathrm{id}_{V_{\bar{x}}} \times m}(\mathrm{id}_{V_{\bar{x}}} \times m)^{-1}(l : [\![\varphi]\!] \to V_{\bar{x}} \times V).$$

Where the maps l, p and p' are as in the diagram above. Since $p' = p \circ (\mathrm{id}_{V_{\bar{x}}} \times m)$ it follows that $\forall_{p'} = \forall_p \forall_{(\mathrm{id}_{V_{\bar{x}}} \times m)}$. This shows (4.5). The proofs for the more general cases are very similar. □

The next proposition summarizes the results announced in the introduction.

Proposition 4.3.3. *Let $m : A \to V$ and $n : B \to M$ be small subobjects of V classified by $a : 1 \to V$ and $b : 1 \to V$, respectively.*

Products:

$$A \times B \cong [\![x : V \mid \exists z, y(z\varepsilon a \wedge y\varepsilon b \wedge x = \langle z, y\rangle)]\!] \text{ in } \mathrm{Sub}(V),$$

where $A \times B$ is regarded as a subobject of V via the map $\langle \text{-}, \text{-}\rangle \circ m \times n$. So if $a \times b : 1 \to V$ classifies $A \times B$, we have

$$(V, \varepsilon) \models x\varepsilon a \times b \leftrightarrow \exists z, y(z\varepsilon a \wedge y\varepsilon b \wedge x = \langle z, y\rangle) \tag{4.6}$$

Powersets:

$$\mathcal{P}_{\mathrm{s}}(A) \cong [\![x : V \mid x \subseteq a]\!] \text{ in } \mathrm{Sub}(V),$$

where $\mathcal{P}_{\mathrm{s}}(A)$ is a subobject of V via the map $\mathrm{Int} \circ \mathcal{P}_{\mathrm{s}}(m)$. So if this subobject is classified by $\mathbf{pow}(a) : 1 \to V$ we have

$$(V, \varepsilon) \models x\varepsilon\, \mathbf{pow}(a) \leftrightarrow x \subseteq a.$$

Functions: *If $f : A \to B$ is an arrow between A and B, then*

$$\Gamma(f) \cong A \xrightarrow{\langle \mathrm{id}, f\rangle} A \times B \xrightarrow{m \times n} V \times V \xrightarrow{\langle \text{-}, \text{-}\rangle} V$$

is a small subobject and suppose this is classified by $\rho : 1 \to V$. Then

$$(V, \varepsilon) \models \text{“}\rho \textit{ is a function from } a \textit{ to } b\text{”}.$$

Proof. • First observe that by the proof of Proposition 2.3.15, $A \times B$ is indeed a small subobject of V. In $\mathrm{Sub}(V)$ we have:

$$\begin{aligned} & A \times B \\ = \;& [\![x \mid \exists a : A \exists b : B(x = \langle m(a), n(b)\rangle)]\!] \\ = \;& [\![x \mid \exists y : V \exists z : V(y\varepsilon a \wedge z\varepsilon b \wedge x = \langle y, z\rangle)]\!], \end{aligned}$$

where for the last equality we use Lemma 4.3.2.

- Let $a' : 1 \to \mathcal{P}_s(V)$ be the classifying map of $m : A \to V$ that we obtain from (P1) so that $\mathrm{Int} \circ a' = a$. We show that $h : [\![x : \mathcal{P}_s(V) \mid x \subseteq_V a']\!] \to \mathcal{P}_s(V)$ and $\mathcal{P}_s(m) : \mathcal{P}_s(A) \to \mathcal{P}_s(V)$ are the same subobjects of $\mathcal{P}_s(V)$. Our claim follows from this as using Lemma 4.1.3, it is easy to see that $\exists_{\mathrm{Int}}[\![x : \mathcal{P}_s(V) \mid x \subseteq_V a']\!] = [\![x : V \mid x \subseteq a]\!]$ in $\mathrm{Sub}(V)$.

  We show two directions separately. The proof of $\mathcal{P}_s(A) \leq [\![x : \mathcal{P}_s(V) \mid x \subseteq_V a']\!]$ is a direct application of the internal logic. Since by earlier results we have:

$$
\begin{array}{lll}
 & \mathcal{P}_s(A) & \\
= & [\![x : \mathcal{P}_s(V) \mid \exists C : \mathcal{P}_s(A). x = \mathcal{P}_s(m)(C)]\!] & \\
= & [\![x : \mathcal{P}_s(V) \mid \exists C : \mathcal{P}_s(A). \forall y : V(y \in_V x \leftrightarrow y \in_V \mathcal{P}_s(m)(C)]\!] & \text{[By typed extensionality.]} \\
= & [\![x : \mathcal{P}_s(V) \mid \exists C : \mathcal{P}_s(A). \forall y : V(y \in_V x \leftrightarrow \exists z : A(z \in_A C \wedge m(z) = y))]\!] & \text{[2.3 in Lemma 2.3.4.]} \\
\leq & [\![x : \mathcal{P}_s(V) \mid \forall y : V(y \in_V x \to \exists z : A(m(z) = y)]\!] & \text{[Arguing internally.]} \\
\leq & [\![x : \mathcal{P}_s(V) \mid \forall y : V\,(y \in_V x \to y \in_V a')]\!] & \\
= & [\![x : \mathcal{P}_s(V) \mid x \subseteq_V a']\!] & .
\end{array}
$$

  Conversely, we show that the following triangle commutes:

$$
\begin{tikzcd}
{[\![x : \mathcal{P}_s(V) \mid x \subseteq_V a']\!]} \arrow[r, tail, "h"] \arrow[dr, tail, "h"'] & \mathcal{P}_s(V) \arrow[r, "m'"] & \mathcal{P}_s(A) \arrow[dl, tail, "\mathcal{P}_s(m)"] \\
 & \mathcal{P}_s(V) &
\end{tikzcd},
$$

  where the map m' results by applying the contravariant powerset functor to the small map m. Recall that we defined this functor in Lemma 2.3.7. We will argue internally to show that the above diagram commutes. First note that we have

$$\mathcal{E} \models \forall y : V(y \in_V h(x) \leftrightarrow y \in_V h(x) \wedge y \in_V a').$$

  Now also

$$
\begin{array}{ll}
\mathcal{E} \models \forall y : V(y \in_V \mathcal{P}_s(m)m'h(x) & \\
\quad \leftrightarrow \exists z : A(m(z) = y \wedge z \in_A m'h(x)) & \text{[See 2.3, Lemma 2.3.4.]} \\
\quad \leftrightarrow \exists z : A(m(z) = y \wedge (\exists y' : V(y' \in_V h(x) \wedge m(z) = y')) & \text{[See Lemma 2.3.7.]} \\
\quad \leftrightarrow \exists z : A(y \in_V h(x) \wedge m(z) = y)) & \\
\quad \leftrightarrow y \in_V h(x) \wedge \exists z : A\, m(z) = y & \\
\quad \leftrightarrow y \in_V h(x) \wedge y \in a'). &
\end{array}
$$

  So by typed extensionality the diagram commutes.

- Assume that ρ is the classifying map of the graph of $f : A \to B$. In order to show that $(V, \varepsilon) \models$ "ρ is a function", we need to check that the following three statements are satisfied.

$$
\begin{array}{l}
(V, \varepsilon) \models \rho \subseteq a \times b \\
(V, \varepsilon) \models \forall v(v \varepsilon a \to \exists w(w \varepsilon b \wedge \langle v, w \rangle \varepsilon \rho) \quad \text{and} \\
(V, \varepsilon) \models \forall w, v, z(\langle z, v \rangle \varepsilon \rho \wedge \langle z, w \rangle \varepsilon \rho \to v = w).
\end{array}
$$

From the description of ρ it is easy to see that

$$\mathcal{E} \models_{v:V} v\varepsilon\rho \leftrightarrow \exists x : A(v = \langle m(x), n(f(x))\rangle). \tag{4.7}$$

Therefore, we can calculate internally:

$$\begin{aligned}
\mathcal{E} \models v\varepsilon\rho &\to \exists x : A(v = \langle m(x), n(f(x))\rangle) && \text{[By (4.7).]}\\
&\to \exists x : A \exists y : B(v = \langle m(x), n(y)\rangle) &&\\
&\to \exists u : V \exists w : V(u\varepsilon a \wedge w\varepsilon b \wedge v = \langle u, w\rangle) && \text{[By Lemma 4.3.2.]}\\
&\to v\varepsilon a \times b && \text{[By 4.6].}
\end{aligned}$$

So in particular, we have that $(V, \varepsilon) \models v\varepsilon\rho \to v\varepsilon a \times b$, which shows the first claim.

$$\begin{aligned}
\mathcal{E} \models u\varepsilon a &\to \exists x : A(m(x) = u) && \text{[By Lemma 4.3.2.]}\\
&\to \exists x : A(m(x) = u \wedge \langle u, n(f(x))\rangle\varepsilon\rho) && \text{[By (4.7).]}\\
&\to \exists y : B\langle u, n(y)\rangle\varepsilon\rho &&\\
&\to \exists w : V(w\varepsilon b \wedge \langle u, w\rangle\varepsilon\rho) && \text{[By Lemma 4.3.2.]}
\end{aligned}$$

which shows the second claim. Finally,

$$\begin{aligned}
\mathcal{E} \models &\langle z, v\rangle\varepsilon\rho \wedge \langle z, w\rangle\varepsilon\rho &&\\
&\to \exists x : A\, \exists x' : A(m(x) = z \wedge m(x') = z \wedge \langle z, v\rangle\varepsilon\rho \wedge \langle z, w\rangle\varepsilon\rho) && \text{[By (4.7).]}\\
&\to \exists x : A\langle m(x), v\rangle\varepsilon\rho \wedge \langle m(x), w\rangle\varepsilon\rho && \text{[Since } m \text{ is a mono.]}\\
&\to \exists x : A\,(v = n(f(x)) = w) && \text{[Use (4.7).]}\\
&\to v = w.
\end{aligned}$$

So we proved that V believes that ρ is a function.

□

As expected, also more specific properties of the "squeezed" objects remain valid in the universe. The next corollary expresses that squeezed monos give rise to injective functions.

Corollary 4.3.4. *If $f : A \to B$ is a mono in $\mathcal{E}$ and $\rho : 1 \to V$ classifies the graph of f as above, then*

$$(V, \varepsilon) \models \text{“}\rho \text{ is an injective function from } a \text{ to } b.\text{”}$$

Proof. Above we saw that V believes that ρ is a function from a to b so it remains to show that V believes that ρ is injective. We calculate internally:

$$\begin{aligned}
\mathcal{E} \models &\langle z, v\rangle\varepsilon\rho \wedge \langle z', v\rangle\varepsilon\rho\\
&\to \exists x : A\, \exists x' : A\, m(x) = z \wedge m(x') = z' \wedge n(f(x)) = v = n(f(x'))\\
&\to \exists x : A\, \exists x' : A\, m(x) = z \wedge m(x') = z' \wedge x = x'\\
&\to z = z',
\end{aligned}$$

where in the second last step we used that f and n are monos. This proves the claim. □

Recall that in Section 2.3.5 we saw that for a small object A and an object B we can form the exponential B^A. Similarly, we also defined the object $\mathbf{Epi}(A, B)$, the object of epimorphisms from A to B. In the remaining part of this section we will see that we can squeeze these constructions in the same manner as we did in the proposition above.

In the following let $m : A \to V$ and $n : B \to V$ be fixed small subobjects of V, classified by maps $a : 1 \to V$ and $b : 1 \to V$, respectively. Note that the composite

$$\mathcal{P}_\mathrm{s}(A \times B) \xrightarrow{\mathcal{P}_\mathrm{s}(m \times n)} \mathcal{P}_\mathrm{s}(V \times V) \xrightarrow{\mathcal{P}_\mathrm{s}(\langle\text{-},\text{-}\rangle)} \mathcal{P}_\mathrm{s}(V) \xrightarrow{\mathrm{Int}} V \tag{4.8}$$

is a small subobject of V. Assume that it is classified by classified by $\mathbf{pow}(a \times b) : 1 \to V$. By Proposition 4.3.3 it is immediate that

$$(V, \varepsilon) \models x \varepsilon\, \mathbf{pow}(a \times b) \leftrightarrow (x \subseteq a \times b). \tag{4.9}$$

Recall that in Section 2.3.5 we defined

$$B^A := [\![f : \mathcal{P}_\mathrm{s}(A \times B) \mid \forall x : A \exists! y : B (x, y) \in_{A \times B} f]\!] \xrightarrow{h} \mathcal{P}_\mathrm{s}(A \times B) \, .$$

Let $l := \mathrm{Int} \circ \mathcal{P}_\mathrm{s}(\langle\text{-},\text{-}\rangle \circ m \times n)$ the composite as in (4.8). Then B^A is a small subobject of V via $l \circ h$. We will soon see that this subobject corresponds to the set of functions from a to b. Before we will proof another little lemma.

Lemma 4.3.5. *Let* $l := \mathrm{Int} \circ \mathcal{P}_\mathrm{s}(\langle\text{-},\text{-}\rangle \circ m \times n)$. *Then*

$$[\![\langle x, y\rangle \in_{A \times B} z]\!] = [\![\langle m(x), n(y)\rangle \varepsilon l(z)]\!] \mathit{in}\ \mathrm{Sub}(A \times B \times \mathcal{P}_\mathrm{s}(A \times B)).$$

Proof. Let $k := \langle\text{-},\text{-}\rangle \circ m \times n$ The right hand side of the equation above is the pullback in

$$\begin{tikzcd} T \arrow[r] \arrow[d] & \varepsilon \arrow[d] \\ A \times B \times \mathcal{P}_\mathrm{s}(A \times B) \arrow[r, "k \times l"] & V \times V \end{tikzcd} .$$

By Lemma 2.3.9 and the definition of ε the two squares in the diagram below are pullbacks

$$\begin{tikzcd} \in_{A \times B} \arrow[r] \arrow[d] & \in_V \arrow[r] \arrow[d] & \varepsilon \arrow[d] \\ A \times B \times \mathcal{P}_\mathrm{s}(A \times B) \arrow[r, "k \times \mathcal{P}_\mathrm{s}(k)"] & V \times \mathcal{P}_\mathrm{s}(V) \arrow[r, "\mathrm{id}_V \times \mathrm{Int}"] & V \times V \end{tikzcd}$$

Therefore, by the two pullback Lemma the outer rectangle is a pullback. Since $k \times l = \mathrm{id}_V \times \mathrm{Int} \circ k \times \mathcal{P}_\mathrm{s}(k)$, it follows that $\in_{A \times B}$ and T define the same subobject of $A \times B \times \mathcal{P}_\mathrm{s}(A \times B)$. This proves the claim. $\square$

Now we are ready to show that $hl : B^A \to V$ gives rise to the subobject of all functions from a to b. Note that since l is mono, the subobject defined by

hl is the same as $\exists_l [\![z : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists ! y : B (x, y) \in_{A \times B} z]\!]$. We calculate in $\mathrm{Sub}(V)$:

$$\begin{array}{ll}
& \exists_l [\![z : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists ! y : B (x, y) \in_{A \times B} z]\!] \\
= & \exists_l [\![z : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists ! y : B \langle m(x), n(y) \rangle \varepsilon l(z)]\!] \quad \text{[By Lemma 4.3.5.]} \\
= & [\![w : V \mid \exists z : \mathcal{P}_{\mathrm{s}}(A \times B) l(z) = w \wedge \forall x : A \exists ! y : B \langle m(x), n(y) \rangle \varepsilon l(z)]\!] \\
= & [\![w : V \mid \exists z : \mathcal{P}_{\mathrm{s}}(A \times B) l(z) = w \wedge \forall x : A \exists ! y : B \langle m(x), n(y) \rangle \varepsilon w]\!] \\
= & [\![w : V \mid w \subseteq a \times b \wedge \forall x : A \exists ! y : B \langle m(x), n(y) \rangle \varepsilon w]\!] \quad \text{[By (4.9).]} \\
= & [\![w : V \mid w \subseteq a \times b \wedge \forall u (u \varepsilon a \to \exists ! v (v \varepsilon b \wedge \langle u, v \rangle \varepsilon w))]\!] \quad \text{[Use Lemma 4.3.3.]}
\end{array}$$

If B^A is classified by the arrow $b^a : 1 \to V$ the above calculation shows that

$$(V, \varepsilon) \models f \varepsilon a^b \leftrightarrow \text{“f is a function from } a \text{ to } b\text{”}.$$

Recall that in Section 2.3.5, we defined

$$\mathbf{Epi}(A, B) := \{ z : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists ! y : B \, (x, y) \in_{A \times B} z \wedge \forall y : B \exists x : A \, (x, y) \in_{A \times B} z \}$$

as a subobject of $\mathcal{P}_{\mathrm{s}}(A \times B)$. Also $\mathbf{Epi}(A, B)$ gives rise to a small subobject of V via the map l. As above we calculate in $\mathrm{Sub}(V)$

$$\begin{array}{ll}
& \exists_l [\![z : \mathcal{P}_{\mathrm{s}}(A \times B) \mid \forall x : A \exists ! y : B (x, y) \in_{A \times B} z \wedge \forall y : B \exists x : A \, (x, y) \in_{A \times B} z]\!] \\
= & [\![w : V \mid w \subseteq a \times b \wedge \forall u (u \varepsilon a \to \exists ! v (v \varepsilon b \wedge \langle u, v \rangle \varepsilon w) \wedge \forall v (v \varepsilon b \to (\exists u (u \varepsilon a \wedge \langle u, v \rangle \varepsilon w))]\!].
\end{array}$$

So if $\mathbf{Epi}(A, B)$ is classified by $\mathrm{epi}(a, b) : 1 \to V$, then

$$(V, \varepsilon) \models f \varepsilon \, \mathrm{epi}(a, b) \leftrightarrow \text{“‘is a surjective function from } a \text{ to } b\text{”}.$$

Corollary 4.3.6. *If* $\mathbf{Epi}(A, B) \cong 0$ *in* $\mathcal{E}$ *then*

$$(V, \varepsilon) \models \textit{“there is no surjective function from } a \textit{ to } b.\textit{”}$$

4.4 Existence of Algebraic Universes

In this section we will discuss the existence of initial algebras for the $\mathcal{P}_{\mathrm{s}}$-functor on a category of classes. First, we will examine a criterion ensuring the existence of such initial algebras. Next, we will discuss whether the criterion is applicable to the examples of categories of classes that we considered in Chapter 3. In particular, we will prove that this is the case for the category Classes and the category of large presheaves. We also briefly discuss the case of the category of large sheaves.

Let $\mathcal{P}_s : \mathcal{E} \to \mathcal{E}$ be an endofunctor on a category $\mathcal{E}$ and let $\mathcal{D}$ be an index category. Assume $\mathcal{F} : \mathcal{D} \to \mathcal{E}$ is a diagram with colimit $\eta : \mathcal{F} \Rightarrow \Delta_Z$. Then $\mathcal{P}_s(\eta) : \mathcal{P}_s \mathcal{F} \Rightarrow \Delta_{\mathcal{P}_s(Z)}$ is a cocone for the diagram $\mathcal{P}_s \circ \mathcal{F} : \mathcal{D} \to \mathcal{E}$. However, the cocone $\mathcal{P}_s(\eta)$ need not to be colimiting. If $\mu : \mathcal{P}_s \mathcal{F} \Rightarrow \Delta_W$ is the colimit for $\mathcal{P}_s \circ \mathcal{F}$, then there is a canonical map

$$k : W \to \mathcal{P}_s(Z)$$

of cocones. We say that the functor $\mathcal{P}_s$ *commutes with colimits of type* $\mathcal{D}$, if k is an isomorphism for every diagram over $\mathcal{D}$.

If the endofunctor $\mathcal{P}_s : \mathcal{E} \to \mathcal{E}$ commutes with colimits of type $(\mathrm{Ord}, \leq)$ it admits an initial algebra as the criterion in Proposition 4.4.1 ensures. In the case that the endofunctor $\mathcal{P}_\mathrm{s} : \mathcal{E} \to \mathcal{E}$ is the powerclass functor on some category of classes, the criterion tells us that an initial algebra for $\mathcal{P}_\mathrm{s}$ can be obtained by mimicking the construction of a cumulative hierarchy inside our category of classes.

Proposition 4.4.1. *Assume that $\mathcal{E}$ has colimits of type $(\mathrm{Ord}, \leq)$. Moreover, assume that $\mathcal{P}_s : \mathcal{E} \to \mathcal{E}$ commutes with colimits of this type. Then $\mathcal{P}_s$ has an initial algebra $\alpha : \mathcal{P}_s(A) \to A$.*

Proof. First we will construct a diagram $\mathcal{F} : (Ord, \leq) \to \mathcal{E}$. Then we show that its colimit gives rise to an initial algebra of $\mathcal{P}_s$. By an external recursion over the ordinals we define:

$\mathcal{P}_s^0(0) := 0$ and $\iota_{0,1} : 0 \to \mathcal{P}_s(0)$ the unique map.

For successor steps:

$\mathcal{P}_s^{\alpha+1}(0) := \mathcal{P}_s(\mathcal{P}_s^{\alpha}(0))$ and $\iota_{\alpha+1,\alpha+2} := \mathcal{P}_s(\iota_{\alpha,\alpha+1})$

For a limit β, let $\eta : (\mathcal{P}_s^{\alpha}(0), \iota_{\alpha,\alpha+1})_{\alpha<\beta} \Rightarrow \Delta_X$ be the colimit of the diagram described by $(\mathcal{P}_s^{\alpha}(0), \iota_{\alpha,\alpha+1})_{\alpha<\beta}$.

Let $\mathcal{P}_s^{\beta}(0) := X$ and $\iota_{\alpha,\beta} = \eta_\alpha$ for $\alpha < \beta$.

It is easy to see that also, $\mathcal{P}_s(\mathcal{P}_s^{\beta}(0))$ gives rise to a cocone for $(\mathcal{P}_s^{\alpha}(0), \iota_{\alpha,\alpha+1})_{\alpha<\beta}$, then we define

$\iota_{\beta,\beta+1} : \mathcal{P}_s^{\beta}(0) \to \mathcal{P}_s^{\beta+1}(0)$ resulting from the universal property of $\mathcal{P}_s^{\beta}(0)$.

The above defines a diagram $\mathcal{F} : (\mathrm{Ord}, \leq) \to \mathcal{E}$ given by $(\mathcal{P}_s^{\alpha}(0), \iota_{\alpha,\beta})_{\alpha<\beta}$. Let $\mu : \mathcal{F} \Rightarrow \Delta_A$ be its colimit. Since by assumption the functor $\mathcal{P}_s$ preserves colimits of type $(\mathrm{Ord}, \leq)$, $\mathcal{P}_s(\mu) : \mathcal{P}_s \Rightarrow \Delta_{\mathcal{P}_s(A)}$ is the colimit for $\mathcal{P}_s \circ \mathcal{F}$. It is easy to see that also $(A, \mu_{\alpha+1})_{\alpha\in\mathrm{Ord}}$ gives rise to a cocone for $\mathcal{P}_s \circ \mathcal{F}$. So there is a unique map $s : \mathcal{P}_s(A) \to A$ such that $\mu_{\alpha+1} = s \circ \mathcal{P}_s(\mu_\alpha)$ for all $\alpha \in \mathrm{Ord}$.

Claim. *$s : \mathcal{P}_s(A) \to A$ is the initial $\mathcal{P}$-algebra.*

Assume $t : \mathcal{P}_s(B) \to B$ is another algebra for $\mathcal{P}_s$. Similar as above, one can show that B gives rise to a cocone $\nu : \mathcal{F} \Rightarrow \Delta_B$ of $\mathcal{F}$ with

$$\nu_{\alpha+1} = t \circ \mathcal{P}_s(\nu_\alpha), \text{ for all } \alpha \in \mathrm{Ord}. \tag{4.10}$$

By the universal property of A there is a map $f : A \to B$ such that $\nu_\alpha = f \circ \mu_\alpha$ for all α. We would like to show that f gives rise to a $\mathcal{P}_s$-algebra homomorphism. Using (4.10) it is clear that $(B, t\circ\mathcal{P}_s(\nu_\alpha))$ gives a cocone for $\mathcal{P}\circ\mathcal{F}$ and $t\circ\mathcal{P}_s(f)$ is the unique map from the colimiting cocone $(\mathcal{P}_s(A), \mathcal{P}_s(\mu_\alpha))$ making all involved triangles commute, i.e. $t \circ \mathcal{P}_s(f)$ is the unique map such that $t \circ \mathcal{P}_s(\nu_\alpha) = (t \circ \mathcal{P}_s(f)) \circ \mathcal{P}_s(\mu_\alpha)$ for all α. Now also

$$t \circ \mathcal{P}_s(\nu_\alpha) = \nu_{\alpha+1} = f \circ \mu_{\alpha+1} = f \circ s \circ \mathcal{P}_s(\mu_\alpha) \text{ for all } \alpha.$$

So by the uniqueness just discussed it follows that $t \circ \mathcal{P}_{\mathrm{s}}(f) = f \circ s$. This shows that f is indeed a homomorphism of $\mathcal{P}_s$ algebras and one can easily see that f is the unique such. This finishes the proof of the claim and therefore we proved the proposition. $\square$

Our goal in the remaining part of this section is to discuss whether that Proposition 4.4.1 applies to the categories of classes that we already encountered. In Chapter 3, we already argued that the categories Classes, $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ and $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ admit class-sized colimits so it remains to show that the respective powerclass functors commute with colomits of type $(\mathrm{Ord}, \leq)$. In the following we will analyze even more generally whether the respective powerclass functors commute with colimts over ∞-filtered categories.

A category $\mathcal{D}$ is called *κ-filtered* if every diagram of size κ has a cocone in $\mathcal{D}$. The category $\mathcal{D}$ is called ∞-filtered, if $\mathcal{D}$ is κ-filtered for every κ. Note that in particular the category $(\mathrm{Ord}, \leq)$ is ∞-filtered.

Recall that a colimit of a diagram in Classes is constructed by first taking the coproduct over the images of the diagram and then quotient by an equivalence relation. If the index category if filtered, these equivalence relations take a particularly simple form. In particular, we do not have to look at general "zig-zags". For a proof see [9, p. 77].

Lemma 4.4.2. *Let $\mathcal{D}$ be an ∞-filtered category and let $\mathcal{F} : \mathcal{D} \to \mathrm{Classes}$ be a diagram. Then the limiting cocone $\eta : \mathcal{F} \Rightarrow \Delta_{\lim \mathcal{F}}$ of $\mathcal{F}$ is of the form $\lim(\mathcal{F}) \cong \coprod_{D_i \in \mathcal{D}} \mathcal{F}(D_i)/\sim$, where*

$$\begin{aligned}(i, x) \sim (j, y) \Leftrightarrow & x \in \mathcal{F}(D_i) \text{ and } y \in \mathcal{F}(D_j) \text{ and there are} \\ & f : D_i \to D_k \text{ and } g : D_j \to D_k \text{ for some } D_k \text{ in } \mathcal{D} \\ & \text{such that } \mathcal{F}(f)(x) = \mathcal{F}(g)(y).\end{aligned}$$

Proposition 4.4.3. *Let $\mathcal{D}$ be an ∞-filtered category. Then $\mathcal{P}_{\mathrm{s}} : \mathrm{Classes} \to \mathrm{Classes}$ commutes with colimits of type $\mathcal{D}$.*

Proof. Let $\mathcal{F} : \mathcal{D} \to \mathrm{Classes}$ a diagram. Let $\beta : \mathcal{F} \Rightarrow \Delta_Z$ is the limiting cocone for $\mathcal{F}$ and assume $\alpha : \mathcal{P}_{\mathrm{s}} \mathcal{F} \Rightarrow \Delta_W$ is the limiting cocone for $\mathcal{P}_{\mathrm{s}} \mathcal{F}$. Let

$$u : W \to \mathcal{P}_{\mathrm{s}}(Z) \text{ be the canonical map of cocones.}$$

We will show that u is an isomorphism.

Claim. *u is a mono.*

By Lemma 4.4.2, we know that $W := \coprod \mathcal{P}_{\mathrm{s}} \mathcal{F}(D_i)/\sim$, where $\sim$ as described above. Let $X, Y \in W$ with $u(X) = u(Y)$. Let $x_0 \in \mathcal{P}_{\mathrm{s}}(\mathcal{F}(D_i))$ and $y_0 \in \mathcal{P}_{\mathrm{s}}(\mathcal{F}(D_j))$ be representatives of the equivalence classes X and Y, respectively. We will show that $x_0 \sim y_0$. By assumption, we have that

$$\mathcal{P}_{\mathrm{s}}(\beta_i)(x_0) = u([x_0]) = u([y_0]) = \mathcal{P}_{\mathrm{s}}(\beta_j)(y_0). \tag{4.11}$$

Define

$$\mathcal{L} := \{(a, b) \in x_0 \times y_0 \mid \beta_i(a) = \beta_j(b)\}.$$

By the description of Z according to Lemma 4.4.2 for all $(a, b) \in \mathcal{L}$ we can pick $D_{(a,b)}$ of $\mathcal{D}$ such that there are $f_{(a,b)} : D_i \to D_{(a,b)}$ and $g_{(a,b)} : D_j \to D_{(a,b)}$ with

$\mathcal{F}(f_{(a,b)})(a) = \mathcal{F}(g_{(a,b)})(b)$. Note that since x_0 and y_0 are sets, also $\mathcal{L}$ is a set. So since $\mathcal{D}$ is ∞-filtered, there is D_k in D and $f : D_i \to D_k$ and $g : D_j \to D_k$ such that for all $(a, b) \in \mathcal{L}$ we have $\mathcal{F}(f)(a) = \mathcal{F}(g)(b)$. We will show that $\mathcal{P}_\mathrm{s}(\mathcal{F}(f))(x_0) = \mathcal{P}_\mathrm{s}(\mathcal{F}(g))(y_0)$, finishing the proof that $x_0 \sim y_0$. To this end, let $d \in \mathcal{P}_\mathrm{s}(\mathcal{F}(f))(x_0)$. Since $\mathcal{P}_\mathrm{s}(\mathcal{F}(f))(x_0) = \mathrm{Im}_{\mathcal{F}(f)}(x_0)$, there is $a \in x_0$ such that $\mathcal{F}(f)(a) = d$. The latter implies that $\beta_i(a) = \beta_k(\mathcal{F}(f)(a))$. By (4.11) there is b in y_0 with $\beta_i(a) = \beta_j(b)$, i.e. $(a, b) \in \mathcal{L}$. By the choice of D_k this implies that $\mathcal{F}(g)(b) = \mathcal{F}(f)(a) = d$ in D_k. We conclude that $d \in \mathcal{P}_\mathrm{s}(\mathcal{F}(g))(y_0)$. The other direction follows by symmetry. This finishes the proof that u is mono.

Claim. *u is epi.*

Let $A \in \mathcal{P}_\mathrm{s}(Z)$, i.e. A is a subset of Z. Let A_0 be a system of representatives of the equivalence classes occurring in A. Moreover, for $a \in A_0$ let D_a in $\mathcal{D}$ such that $a \in \mathcal{F}(D_a)$. By ∞-filteredness of $\mathcal{D}$, let D_k in $\mathcal{D}$, $f_a : D_a \to D_k$ a cocone for the collection of D_a's. Then $A_0'' := \{\mathcal{F}(f_a)(a) \mid a \in A_0\} \in \mathcal{P}_\mathrm{s}(\mathcal{F}(D_k))$ and it is easy to see that $\mathcal{P}_\mathrm{s}(\beta_k)(A_0'') = A$. But then u maps $\alpha_k(A_0'') \in W$ to A. This finishes the proof that u is epi. □

Proposition 4.4.1 and Proposition 4.4.3 imply:

Corollary 4.4.4. *The functor* $\mathcal{P}_\mathrm{s} : \mathrm{Classes} \to \mathrm{Classes}$ *admits an initial algebra.*

Using the pointwise construction of colimits in the category of large presheaves it is not difficult to see that also the powerclass functor on the category of large presheaves commutes with colimits over ∞-filtered categories.

Proposition 4.4.5. *Let $\mathcal{C}$ be a small category. Then the powerclass functor* $\mathcal{P}_\mathrm{s} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ *commutes with colimits over ∞-filtered categories.*

Proof. Let $\mathcal{F} : \mathcal{D} \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ be a diagram, where $\mathcal{D}$ ∞-filtered. Let $Z : \mathcal{C}^{\mathrm{op}} \to \mathrm{Classes}, \beta : \mathcal{F} \Rightarrow \Delta_Z$ be the colimit for $\mathcal{F}$. Also let $W : \mathcal{C}^{\mathrm{op}} \to \mathrm{Classes}$, $\alpha : \mathcal{P}_\mathrm{s}\,\mathcal{F} \Rightarrow \Delta_W$ be the colimit for $\mathcal{P}_\mathrm{s}\,\mathcal{F}$. Let $k : W \Rightarrow \mathcal{P}_\mathrm{s}(Z)$ result from the universal property of W.

Claim. *$k : W \to \mathcal{P}_\mathrm{s}(Z)$ is a mono.*

It suffices to show that k is pointwise a mono in Classes. So let $C \in \mathcal{C}$ be fixed. We will show that $k_C : W(C) \to \mathcal{P}_\mathrm{s}(Z)(C)$ is a mono in Classes. Since colimits in $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ are constructed pointwise, using Lemma 4.4.2, we know that $W(C) \cong \coprod \mathcal{P}_\mathrm{s}(\mathcal{F}(D_i))(C)/\sim$, where

$$\begin{aligned}(i, x_0) \sim (j, y_0) \quad \Leftrightarrow \quad & x_0 \in \mathcal{P}_\mathrm{s}(\mathcal{F}(D_i))(C), y_0 \in \mathcal{P}_\mathrm{s}(\mathcal{F}(D_j))(C) \text{ and there are} \\ & f : D_i \to D_k \text{ and } g : D_j \to D_k \text{ for some } D_k \text{ in } \mathcal{D} \text{ such that} \\ & \mathcal{P}_\mathrm{s}(\mathcal{F}(f))_C(x_0) = \mathcal{P}_\mathrm{s}(\mathcal{F}(g))_C(y_0).\end{aligned}$$

Assume $X, Y \in W(C)$ with $k_C(X) = k_C(Y)$. Let $x_0 \in \mathcal{P}_\mathrm{s}(\mathcal{F}(D_i))(C)$ and $y_0 \in \mathcal{P}_\mathrm{s}(\mathcal{F}(D_j))(C)$ be representatives of the equivalence classes of X and Y, respectively. Note that by our description of the $\mathcal{P}_\mathrm{s}$ functor in Section 3.2, we know that $x_0 \subseteq y_C \times F(D_i)$ is a small subpresheaf and similarly $y_0 \subseteq y_C \times \mathcal{F}(D_j)$ small. For C' in $\mathcal{C}$ we define

$$\mathcal{L}_{C'} := \{((h, a), (l, b)) \in x_0(C') \times y_0(C') \mid h = l \text{ and } (\beta_i)_{C'}(a) = (\beta_j)_{C'}(b)\}.$$

Since colimits in $\text{Classes}^{\mathcal{C}^{\text{op}}}$ are computed pointwise, we know that $\beta_{C'} : \mathcal{F}_{C'} \Rightarrow \Delta_{Z(C')}$ is a colimit in Classes, where $\mathcal{F}_{C'} : \mathcal{D} \to \text{Classes}, D \mapsto \mathcal{F}(D)(C')$ and $(\beta_{C'})_i := (\beta_i)_{C'}$. As in the proof of Proposition 4.4.3, we can pick for all pairs $((h,a),(h,b)) \in \mathcal{L}_{C'}$, an element $D_{h,a,b}$ of $\mathcal{D}$ that is an upper bound of D_i and D_j such that a and b are equated in $D_{h,a,b}$ i.e. there are maps $f_{h,a,b} : D_i \to D_{h,a,b}$ and $g_{h,a,b} : D_j \to D_{h,a,b}$ such that $\mathcal{F}(f_{h,a,b})_{C'}(a) = \mathcal{F}(g_{h,a,b})_{C'}(b)$. By ∞-filteredness there is an upper bound $D^{C'}$ for all the $D_{h,a,b}$ so in particular there are maps $f^{C'} : D_i \to D^{C'}$, $g^{C'} : D_i \to D^{C'}$ such that for all $((h,a),(h,b)) \in \mathcal{L}_{C'}$ we have that $\mathcal{F}(f^{C'})_{C'}(a) = \mathcal{F}(g^{C'})_{C'}(b)$. having found such a $D^{C'}$ for all C' in $\mathcal{C}$, we can use ∞-filteredness again -and the fact that $\mathcal{C}$ is small- to find an D in $\mathcal{D}$ and $f : D_i \to D$ and $g : D_j \to D$ such that for all C' in $\mathcal{C}$ and for all $((h,a),(h,b)) \in \mathcal{L}_{C'}$, $\mathcal{F}(f)_{C'}(a) = \mathcal{F}(g)_{C'}(b)$. We claim that

$$\mathcal{P}_{\text{s}}(\mathcal{F}(f))_C(x_0) = \mathcal{P}_{\text{s}}(\mathcal{F}(g))_C(y_0). \tag{4.12}$$

Let C' in $\mathcal{C}$ and assume $(h,d) \in \mathcal{P}_{\text{s}}(\mathcal{F}(f))_C(x_0)(C')$. Then $(h,d) \in y_C(C') \times \mathcal{F}(D)(C')$ such that there is $a \in \mathcal{F}(D_i)(C')$, $(h,a) \in x_0(C')$ and $\mathcal{F}(f)_{C'}(a) = d$. This implies that $(h, (\beta_i)_{C'}(a)) \in \mathcal{P}_{\text{s}}(\beta_i)_C(x_0)(C')$. Now by assumption, we have that $\mathcal{P}_{\text{s}}(\beta_i)_C(x_0)(C') = k_C([x_0])(C') = k_C([y_0])(C') = \mathcal{P}_{\text{s}}(\beta_j)_C(y_0)(C')$. So there is $(h,b) \in y_0(C')$ with $(\beta_j)_{C'}(b) = (\beta_i)_{C'}(a)$. So the triple $(h,a,b) \in \mathcal{L}_{C'}$ and therefore $\mathcal{F}(f)_{C'}(a) = \mathcal{F}(g)_{C'}(b) = d$. So we established $(h,d) \in \mathcal{P}_{\text{s}}(\mathcal{F}(f))_C(y_0)(C')$. This shows that $\mathcal{P}_{\text{s}}(\mathcal{F}(f))_C(x_0)(C') \subseteq \mathcal{P}_{\text{s}}(\mathcal{F}(f))_C(y_0)(C')$. The other direction follows by symmetry and therefore we proved (4.12). But (4.12) implies that $x_0 \sim y_0$ which finishes the proof that k is mono.

Claim. *k is epi.*

First note that it is enough to show that k is componentwise an epi. So let C in $\mathcal{C}$ and let $X \in \mathcal{P}_{\text{s}}(Z)(C)$, i.e. $X \subseteq y_C \times Z$ a small subpresheaf. Our aim is to show that there is a D_k in $\mathcal{D}$ and $Y \subseteq y_C \times \mathcal{F}(D_k)$ a small subpresheaf such that $\mathcal{P}_{\text{s}}(\beta_k)_C(Y) = X$. Then we would be done as $(\alpha_k)_C(Y)$ would be a preimage of X. By ∞-filteredness, there is a D_k such that every equivalence classes occurring in X has a representative $(h,x) \in y_C(C') \times \mathcal{F}(D_k)(C')$. For $C' \in \mathcal{C}$ let $x_0(C') \subseteq y_C(C') \times \mathcal{F}(D_k)(C')$ be a system of representatives of the equivalence classes. Note that $x_0(C')$ is a set for every C' since X is small. Now the collection of the $x_0(C')$ does not necessarily define a subpresheaf of $y_C \times \mathcal{F}(D_k)$, however there is a smallest subpresheaf Y of $y_C \times \mathcal{F}(D_k)$ that such that $x_0(C') \subseteq Y(C')$ for all C' in $\mathcal{C}$. Since all the $x_0(C')$ are sets is clear that Y is small. One can easily calculate that Y is as desired. $\square$

Using Proposition 4.4.1 and Proposition 4.4.5 we can infer:

Corollary 4.4.6. *The functor $\mathcal{P}_{\text{s}} : \text{Classes}^{\mathcal{C}^{\text{op}}} \to \text{Classes}^{\mathcal{C}^{\text{op}}}$ admits an initial algebra.*

We also expect, that the powerclass functor $\mathcal{P}_{\text{s}}^{\text{Sh}} : \text{LSh}(\mathcal{C}, \text{Cov}) \to \text{LSh}(\mathcal{C}, \text{Cov})$ commutes with colimits over ∞-filtered categories, establishing the result that also $\mathcal{P}_{\text{s}}^{\text{Sh}}$ admits an initial algebra. Due to a lack of time we cannot provide a complete proof of this fact. We we still sketch the steps one would need to take. Assume $\mathcal{F} : \mathcal{D} \to \text{LSh}(\mathcal{C}, \text{Cov})$ is a diagram over an ∞-filtered category $\mathcal{D}$. In Chapter 3, we saw that colimits in $\text{LSh}(\mathcal{C}, \text{Cov})$ are obtained by first taking the colimit X of $\mathcal{F}$ in $\text{Classes}^{\mathcal{C}^{\text{op}}}$ and then apply the sheafification functor. However, we expect that the particular form of an ∞-filtered category $\mathcal{D}$ ensures that the colimiting presheaf X is already a sheaf, i.e. we expect that one can prove:

Lemma 4.4.7. $\mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ *is closed under colimits over ∞-filtered categories in* $\mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$.

Note that the powerclass functor $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}} : \mathrm{LSh}(\mathcal{C}, \mathrm{Cov}) \to \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ is in fact already defined on presheaves, i.e. it can be seen as a functor $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$. We expect that one can prove:

Lemma 4.4.8. *The functor $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}} : \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}} \to \mathrm{Classes}^{\mathcal{C}^{\mathrm{op}}}$ commutes with colimits over ∞-filtered categories.*

The proof would be similar to the proof of Proposition 4.4.5, however, one would need to take the closure operator into account that is involved in the definition of $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}$. Lemma 4.4.7 and Lemma 4.4.8 together would yield that $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}} : \mathrm{LSh}(\mathcal{C}, \mathrm{Cov}) \to \mathrm{LSh}(\mathcal{C}, \mathrm{Cov})$ commutes with colimit over ∞-filtered categories.

Chapter 5

An Algebraic Model for $\neg$ CH

Towards the end of the 19th century, Georg Cantor proved by his famous diagonal argument that the real numbers have a strictly larger cardinality than the natural numbers. Cantor conjectured that all subsets of the reals are either countable or have the same cardinality as the reals themselves. The statement became known as the Continuum hypothesis (**CH**) and was considered as so important that Hilbert placed it on top of his list of open mathematical problems to be solved in the 20th century. A partial answer to the conjecture was found in 1940 by Kurt Gödel who proved that **CH** is consistent with the other axioms of Zermelo-Fraenkel by constructing L, the universe of constructible sets [15]. A full answer to the conjecture was found in 1963 by Paul Cohen [11], who introduced the method of forcing to construct a model of set theory falsifying **CH**. The results of Gödel and Cohen established the independence of the Continuum hypothesis from the other axioms of Zermelo-Fraenkel set theory.

Led by the idea that topos theory would serve as an alternative mathematical foundation, people saw themselves challenged to provide a topos theoretic proof of Cohen's result. In 1972 Tierney succeeded to solve this task. In [29], Tierney proved that the category of sheaves over the Cohen poset provides a topos that falsifies the continuum hypothesis and satisfies the axioms of a category of sets (**CS**). In this way, Tierney showed that the continuum hypothesis is not implied by the other axioms for a category of sets. The axioms of a category of sets are essentially the topos axioms including a natural number object, Booleanness and the axiom of choice. They differ from the **ZFC** axioms mostly in the lack of replacement and (unrestricted) separation.

The goal of this chapter is to show that Tierney's proof translates to the context of algebraic set theory. In this way we show that the categorical proof can provide a model for the full **ZFC** axioms -as opposed to only the axiom of **CS**- falsifying the continuum hypotheses. We will proceed as follows: In the first part, we will see that Tierney's result translates to a category of classes by considering the category of large sheaves over the Cohen poset. To this end we will mostly follow the presentation given in Chapter 6 of Mac Lane and Moerdijk's book[20], modifying some arguments in order to make them fit into our framework. In this way, we obtain that the outer category of classes falsifies the continuum hypothesis. In a second step, we will show that the statement in the outer categy of classes translates to the algebaric universe that the category of classes contains. For this purpose we will essentially use the squeezing results that we established in Chapter 3. In this way we will obtain an algebraic universe of set theory falsifying the continuum hypotheses.

In a nutshell, the topos theoretic proof works as follows. One starts with a large set B whose cardinality is even larger than that of the powerset of the natural numbers. The Cohen poset $\mathbb{P}$ is defined to be the set of finite approximations of a function $B \times \mathbf{N} \to 2$ ordered by reverse inclusion. Now the category of sheaves over the Cohen poset with respect to the dense topology forms a Boolean topos satisfying the axiom of choice. This topos is shown to refute **CH** as one can find a particular sheaf whose cardinality is strictly in between the one of the natural numbers sheaf and its powersheaf. The sheaf in question is obtained as the shefification of the constant presheaf taking the value of the powerset of the true natural numbers. It can be seen as the fake powersheaf of the natural numbers.

First, we will describe the Cohen poset. In the real category of classes let $B := \mathcal{P}\,\mathcal{P}(\mathbf{N})$ be the powerset of the powerset of the natural numbers. So in Classes the cardinal inequalities

$$|\,\mathbf{N}\,| < |2^{\mathbf{N}}| < |B| \quad \text{hold.}$$

Let P be the set of finite approximations of functions from $B \times \mathbf{N}$ to 2, i.e

$$p \in P \Leftrightarrow p = \{((b_1, n_1), i_1), \ldots, ((b_k, n_k), i_k)\}$$

with $b_j \in B, n_j \in \mathbf{N}, i_j \in \{0, 1\}$ for all j with $1 \leq j \leq k$ and such that p is a function.

For $p, q \in P$ let

$$q \leq p \Leftrightarrow q \text{ extends } p \text{ as a function}$$

This clearly defines a partial order on P. We will call $\mathbb{P} := (P, \leq)$ the *Cohen poset*. In the following we will work with the category

$$\mathrm{LSh}(\mathbb{P}, \neg\neg),$$

the category of large sheaves over $\mathbb{P}$ with respect to the dense topology that we introduced in Section 3.3.5. In 3.3.4, we proved that $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ is a category of classes. Assume that $\mathrm{V}_{\mathbb{P}}$ is the initial algebra for the powerclass functor $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}$ on $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ with element relation $\varepsilon_{\mathbb{P}}$. By Proposition 3.3.23 and Corollary 4.2.3 it follows that

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \mathbf{ZF}\,.$$

Our goal is to show that

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \neg\,\mathbf{CH}\,.$$

Let us briefly indicate how we will achieve the result above. In Proposition 4.2.2 we saw that $\omega : 1 \to \mathrm{V}_{\mathbb{P}}$, the classifying map of the natural number object in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ determines the natural numbers in $\mathrm{V}_{\mathbb{P}}$. Moreover, in Proposition 4.3.3 we saw that $\mathbf{pow}(\omega) : 1 \to \mathrm{V}_{\mathbb{P}}$, the classifying map of the powerset of the natural numbers determines the powerset of the natural numbers in $\mathrm{V}_{\mathbb{P}}$. We will see that there is a map $x : 1 \to \mathrm{V}_{\mathbb{P}}$ such that:

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \text{“There are injections } m_1 : \omega \to x \text{ and } m_2 : x \to \mathbf{pow}(\omega)\text{”},$$

and

$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models$ "there is no surjection from ω to x and no surjection from x to $\mathbf{pow}(\omega)$".

This would establish that the algebraic universe $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}})$ believes that there is some cardinality between the one of the natural numbers and the cardinality of its powerset, showing that $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}})$ falsifies the continuum hypothesis . In order to prove this result about the universe, we will first establish the result in the outer category of classes.

For a class X, we will denote by $\hat{X}$ the image under the composite

$$\text{Classes} \xrightarrow{\Delta} \mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}} \xrightarrow{\mathbf{a}} \mathrm{LSh}(\mathbb{P}, \neg\neg) \, .$$

where Δ denotes the constant functor that maps a class X to the constant presheaf with value X and $\mathbf{a}$ denotes the sheafification functor as in Section 3.3.2. The remaining part of this section is devoted to establish the following results in the category $\mathrm{LSh}(\mathbb{P}, \neg\neg)$:

There are monos m_1, m_2 as in

$$\widehat{\mathbf{N}} \overset{m_1}{\rightarrowtail} \widehat{\mathcal{P}(\mathbf{N})} \overset{m_2}{\rightarrowtail} \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\widehat{\mathbf{N}}) \, . \tag{5.1}$$

Moreover,

$$\mathbf{Epi}(\widehat{\mathbf{N}}, \widehat{\mathcal{P}(\mathbf{N})}) \cong 0 \tag{5.2}$$

and

$$\mathbf{Epi}(\widehat{\mathcal{P}(\mathbf{N})}, \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\widehat{\mathbf{N}})) \cong 0 \tag{5.3}$$

Note that $\widehat{\mathcal{P}(\mathbf{N})}$ denotes the sheafification of the constant presheaf taking the value of the powerset of the natural numbers in Classes, whereas $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\widehat{\mathbf{N}})$ is the powersheaf of the natural numbers sheaf in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$. We begin by the construction of the monos from (5.1). The existence of m_1 is obvious. In Classes there is a mono $j_1 : \mathbf{N} \to \mathcal{P}_{\mathrm{s}}(\mathbf{N})$ and since Δ and $\mathbf{a}$ preserve monos, $m_1 := \widehat{j_1}$ is a mono in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ as desired. The construction of m_2 is a bit more involved. In Proposition 5.0.11, we will show that there is a mono

$$m_2' : \hat{B} \longrightarrow \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}}) \text{ in } \mathrm{LSh}(\mathbb{P}, \neg\neg), \text{ where as before } B = \mathcal{P}\,\mathcal{P}(\mathbf{N}).$$

Once such a mono m_2' is found, the existence of m_2 is clear. As above, there is mono $j_2 : \mathcal{P}(\mathbf{N}) \to B$ in the category Classes giving rise to a mono $\hat{j_2} : \widehat{\mathcal{P}(\mathbf{N})} \to \hat{B}$ in the category $\mathrm{LSh}(\mathbb{P}, \neg\neg)$. Then we can define m_2 as the composite $m_2' \circ \hat{j_2}$. In the following we make use of the simplified notations that we introduced in Section 3.3.6.

We begin with the construction of m_2'. First we work in the category of large presheaves over $\mathbb{P}$. Let $\mathcal{P}_{\mathrm{s}} : \mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}} \to \mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$ denote the powerclass functor for the category $\mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$.

In the category of large presheaves over $\mathbb{P}$, i.e the category $\mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$ we would like to define a natural transformation

$$f : \Delta B \Longrightarrow \mathcal{P}_{\mathrm{s}}(\Delta\,\mathbf{N})$$

with components

$$\begin{aligned} f_p : &B \longrightarrow \mathcal{P}_{\mathrm{s}}(\Delta\,\mathbf{N})(p) \\ &b \mapsto Z_{pb} \end{aligned}$$

where $\mathcal{P}_{\mathcal{S}}$ denotes the powerclass functor in $\mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$. For $p \in \mathbb{P}$ and $b \in B$ we define

$$\begin{aligned} Z_{pb} : \mathbb{P}^{\mathrm{op}} &\longrightarrow \text{Classes} \text{ by the assignment} \\ q &\mapsto \{n \mid q \leq p \text{ and } q(b,n) = 0\}. \end{aligned}$$

Claim. *The assignment Z_{pb} gives rise to a subpresheaf of $y_p \times \Delta(\mathbf{N})$.*

Proof. First observe that it is clear that for every $q \in \mathbb{P}$, $Z_{pb}(q) \subseteq (y_p \times \Delta(\mathbf{N}))\,(q)$. In order to prove the claim it remains to check that for $q' \leq q$, $Z_{pb}(q) \subseteq Z_{pb}(q')$. To this end let $q' \leq q$ and assume $n \in Z_{pb}(q)$, i.e. $q \leq p$ and $q(b,n) = 0$. But since $q' \leq q$, by transitivity $q' \leq p$. Also, since $q(b,n) = 0$ and q' extends q, it follows that $q'(b,n) = 0$. But this shows that $n \in Z_{pb}(q')$ as desired. This proves the claim. $\square$

The above claim shows that the assignment f_p is well defined. The naturality of the family f is obvious.

Lemma 5.0.9. *The map f defines a mono in $\mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$.*

Proof. It suffices to show that f is componentwise a mono so let $p \in \mathbb{P}$ be fixed. In order to show that $f_p : B \longrightarrow \mathcal{P}_{\mathrm{s}}(\Delta\,\mathbf{N})(p)$ is a mono in Classes, assume that $b, b' \in B$ with $b \neq b'$. We need to show that $f_p(b) \neq f_p(b')$,i.e. $Z_{pb} \neq Z_{pb'}$. To this end let $n \in \mathbf{N}$ be large enough such that

$$(b,n) \notin \mathbf{dom}(p) \text{ and } (b',n) \notin \mathbf{dom}(p).$$

Let $q := p \cup \{((b,n),0)\} \cup \{((b,n),1)\}$. One can easily see that $n \in Z_{pb}(q)$ but $n \notin Z_{pb'}(q)$. Therefore, $Z_{pb} \neq Z_{pb'}$ as desired. $\square$

Lemma 5.0.10. *Z_{pb} is a closed subpresheaf of $y_p \times \Delta(\mathbf{N})$.*

Proof. Let $q \in \mathbb{P}$. Assume $n \in y_p(q) \times \mathbf{N}$ and assume that the set $D := \{d \mid n \in Z_{pb}(d)\}$ is covering in the dense topology, i.e. D is dense below q. We need to show that $n \in Z_{pb}(q)$. Since $n \in y_p(q) \times \mathbf{N}$ it is clear that $q \leq p$. Next, we argue that $(b,n) \in \mathbf{dom}(q)$. For if not, then $q' := q \cup \{((b,n),1)\} \leq q$ in $\mathbb{P}$. Since D is dense below q there is $d \in D$ with $d \leq q'$. Now $d \in D$ implies that $n \in Z_{pb}(d)$ which implies that $d(b,n) = 0$ but this contradicts that d extends q'. We conclude that $(b,n) \in \mathbf{dom}(q)$. With very similar arguments we can show that $q(b,n) = 0$. But this shows that $n \in Z_{pb}(q)$ and so Z_{pb} is indeed a closed subpresheaf of $y_p \times \Delta(\mathbf{N})$. $\square$

Lemma 5.0.11. *The arrow f gives rise to a mono $m_2' : \hat{B} \Longrightarrow \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\mathbf{a}(\Delta\,\mathbf{N}))$ in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$.*

Proof. By the previous Lemma and the explicit description of the powerclass functor for large sheaves in Proposition 3.3.19, f factors through $\iota \circ \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta(\mathbf{N}))$, where ι is the inclusion $\mathrm{LSh}(\mathbb{P}, \neg\neg) \to \mathbf{Classes}^{\mathbb{P}^{\mathrm{op}}}$. Since $\mathbf{a}$ preserves monos the map

$$\mathbf{a}(f) : \hat{B} \Longrightarrow \mathbf{a}\,\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta\,\mathbf{N}) \text{ is mono in } \mathrm{LSh}(\mathbb{P}, \neg\neg). \tag{5.4}$$

Since $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta(\mathbf{N}))$ is a sheaf, it follows that $\mathbf{a}\,\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta\,\mathbf{N}) \cong \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta(\mathbf{N}))$. As in Corollary 5 in [20, p. 282], one can also show that $\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta(\mathbf{N})) \cong \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\mathbf{a}(\Delta(\mathbf{N})))$. The two isomorphisms together give $\mathbf{a}\,\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\Delta\,\mathbf{N}) \cong \mathcal{P}_{\mathrm{s}}(\mathbf{a}(\Delta(\mathbf{N})))$. Therefore, the claim follows from (5.4). $\square$

With Lemma 5.0.11, we showed that we have monos

$$\widehat{\mathbf{N}} \overset{m_1}{\rightarrowtail} \widehat{\mathcal{P}(\mathbf{N})} \overset{m_2}{\rightarrowtail} \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\widehat{\mathbf{N}})$$

as stated in (5.1). Our next goals are to establish (5.2) and (5.3) above, i.e. to show that in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$

$$\mathbf{Epi}(\widehat{\mathbf{N}}, \widehat{\mathcal{P}(\mathbf{N})}) \cong 0 \text{ and } \mathbf{Epi}(\widehat{\mathcal{P}(\mathbf{N})}, \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\widehat{\mathbf{N}})) \cong 0.$$

The result that will essentially help us to prove the above is stated in Proposition 5.0.13 below. Namely, that for two infinite sets S, T of Classes we have

$$\mathbf{Epi}(S, T) \cong 0 \text{ in Classes implies } \mathbf{Epi}(\widehat{S}, \widehat{T}) \cong 0 \text{ in } \mathrm{LSh}(\mathbb{P}, \neg\neg). \tag{5.5}$$

This is a special property of $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ that follows from a combinatorial property of $\mathbb{P}$ called the countable chain condition. Since $\mathbf{a} \circ \Delta$ always preserves monos, the entailment (5.5) expresses that the functor $\mathbf{a} \circ \Delta$ preserves cardinalities.

Definition 5.0.12. An object X of $\mathcal{E}$ satisfies the *countable chain condition* if in the Heyting algebra $\mathrm{Sub}(X)$ every antichain is countable, i.e. for every collection $\{U_i\}_{i \in I}$ of subobjects of X such that $U_i \wedge U_j = 0$ in $\mathrm{Sub}(X)$ if $i \neq j$, the index set I is necessarily countable. A category of classes $\mathcal{E}$ satisfies the *countable chain condition* if it is generated by objects having the countable chain condition.

Proposition 5.0.13. *If $\mathcal{E}$ satisfies the countable chain condition, then for any two infinite sets S and T in* Classes*,* $\mathbf{Epi}(S, T) \cong 0$ *in* Classes *implies that* $\mathbf{Epi}(\widehat{S}, \widehat{T}) \cong 0$ *in* $\mathrm{LSh}(\mathbb{P}, \neg\neg)$.

Proof. We sketch the proof of Proposition 6 in [20]. Let S and T be infinite sets and suppose that $\mathbf{Epi}(\hat{S}, \hat{T}) \ncong 0$ in $\mathcal{E}$. By our assumption on the generators, it follows that there is an object $E \ncong 0$ in $\mathcal{E}$ satisfying the countable chain condition and an arrow $f : E \to \mathbf{Epi}(\hat{S}, \hat{T})$. In the proof of Proposition 2.3.18, we showed that f corresponds to an epi

$$\begin{tikzcd} E \times \hat{S} \arrow[rr, "g"] \arrow[dr] & & E \times \hat{T} \arrow[dl] \\ & E & \end{tikzcd}$$

in $\mathcal{E}/E$. One can easily show that $g : E \times \hat{S} \to E \times \hat{T}$ is also epi in $\mathcal{E}$. For two elements $s : 1 \to S$ and $t : 1 \to T$ let $\hat{s} : 1 \to \hat{S}$ and $\hat{t} : 1 \to \hat{T}$ be the corresponding global sections in $\mathcal{E}$. Consider the following diagram in $\mathcal{E}$, where both squares are pullbacks.

$$\begin{tikzcd}
U_{s,t} \arrow[r, tail] \arrow[d, tail, "n'"] & P_t \arrow[r, two heads, "h"] \arrow[d, tail, "n"] & E \times 1 \cong E \arrow[d, tail, "1 \times \hat{t}"] \\
E \cong E \times 1 \arrow[r, tail, "1 \times \hat{s}"] & E \times \hat{S} \arrow[r, two heads, "g"] & E \times \hat{T}
\end{tikzcd}$$

Observe that internally,

$$P_t = [\![e : E \, s : \hat{S} \mid \pi_2 g(e, s) = \hat{t}]\!] \text{ in } \mathrm{Sub}(E \times \hat{S})$$

and

$$U_{s,t} = [\![e : E \mid \pi_2 g(e, \hat{s}) = \hat{t}]\!] \text{ in } \mathrm{Sub}(E).$$

In Classes let $W := \{(s, t) \in S \times T \mid U_{s,t} \ncong 0 \text{ in } \mathcal{E}\}$. The next step is to show that for any $t \in T$ there is some $s \in S$ such that $(s, t) \in W$. Following [20], there is an iso

$$\amalg_{s \in S} U_{s,t} \cong P_t \text{ in } \mathcal{E}.$$

Since h is the pullback of an epi, h itself is an epi. Since $E \ncong 0$ it follows that there is at least one $s \in S$ with $U_{s,t} \ncong 0$. This shows that the projection

$$\begin{aligned} W &\longrightarrow T \\ (s, t) &\mapsto t \end{aligned}$$

is epi in Classes. Again, as in [20], one can show that for distinct t, t' in T, $U_{s,t}$ and $U_{s,t'}$ define disjoint subobjects of E. Since E satisfies by assumption the countable chain condition, it follows that for every s there can be at most countably many t such that $U_{s,t} \ncong 0$ so the set $W_s := \{t \in T \mid (s, t) \in W\}$ is countable for every $s \in S$. Since S is infinite by assumption, it follows that the cardinality of $W = \bigcup_{s \in S} W_s$ equals the cardinality of S. The bijection $S \to W$ composed with the epi $W \to T$ from above gives an epi $S \to T$. This shows that $\mathbf{Epi}(S, T) \ncong 0$ in Classes. This finishes the proof of the proposition. □

It remains to show that $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ satisfies the countable chain condition. We will infer this from the well known fact that the Cohen poset satisfies the countable chain condition, i.e. for every family $A \subseteq \mathbb{P}$ such that any two distinct elements of A are incompatible is necessarily countable. The proof of the latter relies on a combinatorial fact called the Δ-system lemma. See for example [19, p. 204].

Lemma 5.0.14. *The Cohen poset $\mathbb{P}$ satisfies the countable chain condition.*

Lemma 5.0.15. $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ *satisfies the countable chain condition.*

Proof. The proof can be found in Lemma 7 [20, p. 289]. We will summarize the main arguments. First, we show that the representable presheaves on $\mathbb{P}$ are in fact sheaves. So let y_p be a the representable functor for p. We will show for any q in $\mathbb{P}$,

$$\text{If } D \text{ is dense below } q \text{ and all elements of } D \text{ are smaller than } p, \text{ then } q \leq p. \tag{5.6}$$

It is not difficult to see that (5.6) implies that y_p is a sheaf. So let D be dense below q and assume that for every $d \in D$, $d \leq p$. We want to show that $q \leq p$. First we show $\mathbf{dom}\, p \subseteq \mathbf{dom}\, q$. For if not, let (b, n) in $\mathbf{dom}(p)$ with $(b, n) \notin \mathbf{dom}(q)$. Without loss of generality assume that $p((b, n)) = 0$. Then $q' := q \cup \{(b, n), 1)\} \leq q$. Since D is dense below q, there is $d \in D$ with $d \leq q'$ so in particular $d((b, n)) = 1$. But this contradicts $d \leq p$. This shows that $\mathbf{dom}(p) \subseteq \mathbf{dom}(q)$. Assume $(b, n) \in \mathbf{dom}(p)$ and without loss of generality $p((b, n)) = 0$. Then also $q((b, n)) = 0$ since otherwise q and b cannot have any common predecessor. This proves (5.6).

Recall from Proposition 3.2.4 that the collection of $\mathbf{a}(y_p)$ generate $\mathrm{LSh}(\mathbb{P}, \neg\neg)$. Since we just showed that all representable presheaves are sheaves, it follows that $\mathbf{a}(y_p) \cong y_p$ for every $p \in \mathbb{P}$ and therefore $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ is generated by the representables. By our considerations in Section 3.3.6 it follows that the representables are subobjects of the terminal object 1. So it suffices to show that 1 satisfies the countable chain condition since this property is clearly inherited by subobjects. To this end let $(U_i)_{i \in I}$ be an antichain in Sub(1). Since the representables generate, for each $i \in I$ we can pick $p_i \in \mathbb{P}$ with $y_{p_i} \leq U_i$. Clearly, also the y_{p_i} form an antichain. But this implies that the family p_i is an antichain in $\mathbb{P}$. By Lemma 5.0.15 the family $(p_i)_{i \in I}$ must be countable. But this proves the claim. □

We will now summarize the stated results and see how we can infer (5.2) and (5.3) from above. By Lemma 5.0.15, $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ satisfies the countable chain condition. So we can apply Proposition 5.0.13 to obtain that

$$\mathbf{Epi}(\widehat{\mathbf{N}}, \widehat{\mathcal{P}(\mathbf{N})}) \cong 0 \text{ in } \mathrm{LSh}(\mathbb{P}, \neg\neg)$$

this establishes (5.2). Again, by applying Lemma 5.0.13 and the fact that B is of larger cardinality than the powerset of the natural numbers in Classes, we obtain

$$\mathbf{Epi}(\widehat{\mathcal{P}(\mathbf{N})}, \hat{B})) \cong 0 \text{ in } \mathrm{LSh}(\mathbb{P}, \neg\neg)\ .$$

Now in (5.0.11), we showed that there is a mono $\hat{B} \to \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}})$ in $\mathcal{E}$. Since $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ is Boolean by Proposition 3.3.23, and clearly there is some global section $1 \to \hat{B}$, we can apply Lemma 2.3.19 to conclude that

$$\mathbf{Epi}(\widehat{\mathcal{P}(\mathbf{N})}, \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}}))) \cong 0 \text{ in } \mathrm{LSh}(\mathbb{P}, \neg\neg)\ .$$

This shows (5.3).

Now we will show that the universe in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ provides an algebraic model of set theory that satisfies the negation of the continuum hypothesis. Having the squeezing results from Section 4.3 at hand, it is straightforward to infer that the universe $\mathrm{V}_{\mathbb{P}}$ falsifies $\mathbf{CH}$ from Tierney's result. So let $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}})$ be the initial algebra for the powerclassfunctor in $\mathrm{LSh}(\mathbb{P}, \neg\neg)$.

Theorem 5.0.16. $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \mathbf{ZFC} + \neg\,\mathbf{CH}$.

Proof. Since $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ is a Boolean pretopos by Corollary 3.3.25, it follows from Corollary 4.2.3 that $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \mathbf{ZF}$. By our discussion is Section 4.2.1 we can also infer that the axiom of choice is validated in $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}})$. Observe in the proof of Theorem 4.2.2 we saw that the natural number object $\hat{\mathbf{N}}$ of $\mathrm{LSh}(\mathbb{P}, \neg\neg)$ defines a small subobject $n : \hat{\mathbf{N}} \to \mathrm{V}_{\mathbb{P}}$ of the universe that gives rise to an arrow $\omega : 1 \to \mathrm{V}_{\mathbb{P}}$ such that

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \text{“}\omega \text{ is the set of natural numbers”}.$$

In Section 4.3, we saw that composite

$$\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}}) \xrightarrow{\mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(n)} \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\mathrm{V}_{\mathbb{P}}) \xrightarrow{\mathrm{Int}} \mathrm{V}_{\mathbb{P}}$$

is a small subobject whose classifying map $\mathbf{pow}(\omega) : 1 \to V$ determines the powerset of ω in $\mathrm{V}_{\mathbb{P}}$, i.e.

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \text{ “}\mathbf{pow}(\omega) \text{ is the powerset of } \omega\text{”}.$$

We will now show that

$$\begin{aligned}(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \exists x\, (\, &\text{“there are injections } m_1 : \omega \to x \wedge m_2 : x \to \mathbf{pow}(\omega) \\ &\wedge \text{ there is no surjection from } \omega \text{ to } x \\ &\wedge \text{ there is no surjection from } x \text{ to } \mathbf{pow}(\omega)\text{”}).\end{aligned}$$

Let $X := \widehat{\mathcal{P}(\mathbf{N})}$, where as previously $\mathcal{P}(\mathbf{N})$ denotes the powerclass of the natural numbers in Classes. By (5.1) from the previous section, there is a mono $m_2 : X \to \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}})$ which gives rise to a small subobjet of $\mathrm{V}_{\mathbb{P}}$ via $\mathrm{Int} \circ \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(n) \circ m_2$. Let this subobject be classified by the map $x : 1 \to \mathrm{V}_{\mathbb{P}}$. In the previous section we showed that there are monos $m_1 : \hat{\mathbf{N}} \to X$ and $m_2 : X \to \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}})$.

In Section 4.3 we examined how we can squeeze these monos into the universe. So if $\rho_1 : 1 \to \mathrm{V}_{\mathbb{P}}$ and $\rho_2 : 1 \to \mathrm{V}_{\mathbb{P}}$ classify the monos m_1 and m_2, respectively by Corollary 4.3.4 we have that

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \text{“}\rho_1 : \omega \to x \text{ and } \rho_2 : x \to \mathbf{pow}(\omega) \text{ are injections”}.$$

According to (5.2) and (5.3) from the previous section we have

$$\mathbf{Epi}(\hat{\mathbf{N}}, X) \cong 0 \text{ and } \mathbf{Epi}(X, \mathcal{P}_{\mathrm{s}}^{\mathrm{Sh}}(\hat{\mathbf{N}})) \cong 0.$$

By Corollary 4.3.6, it follows that

$$(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \text{“there is no surjection } \omega \to x\text{”} \wedge \text{“there is no surjection } x \to \mathbf{pow}(\omega)\text{.”}$$

This shows that the universe $\mathrm{V}_{\mathbb{P}}$ believes that $x : 1 \to \mathrm{V}_{\mathbb{P}}$ has a cardinality in between the one of the natural numbers and the one of the continuum. This shows that $(\mathrm{V}_{\mathbb{P}}, \varepsilon_{\mathbb{P}}) \models \neg\,\mathbf{CH}$. This finishes the proof of the theorem. $\square$

Chapter 6

Conclusion

6.1 Summary of the Thesis

In this thesis we aimed to provide an introduction to the field of the algebraic set theory with the aim on translating Tierney's topos theoretic forcing proof in the setting of algebraic set theory.

In the first chapter, we examined the framework of algebraic set theory in general. To this end we discussed the notion of a Heyting pretopos, gave an introduction to categorical logic and in particular discussed internal languages of categories. We provided an axiomatization of the class of small maps leading to the definition of a category of classes. Moreover, we collected basic properties of categories of classes. The results presented in this chapter were collected mostly from the works [18, 28, 30, 26].

Next, we discussed how categories of presheaves and sheaves on a site fit into the framework of algebraic set theory. That presheaves and sheaves fit into the framework of algebraic set theory was already observed in the original work [18]. Categories of presheaves and sheaves consist of contravariant functors from a small category $\mathcal{C}$ to the category of sets. We saw that such functor categories can be turned into categories of classes by allowing the functors to have classes in their codomain and declaring a map between to such functors, i.e. a natural transformation, as small if all its components are small maps in the category of the true classes. We discussed how the well known topos theoretic proofs also apply in the algebraic set theory setting. In particular, we provided an explicit description of the powerclass functor for the case of large sheaves on a site.

In the next chapter, we focused on the universes of set theory that algebraic set theory provides. These universes arise as initial algebras for an endofunctor $\mathcal{P}_{\mathrm{s}} : \mathcal{E} \to \mathcal{E}$ on a category of classes. We explained how the language of set theory is interpreted in these algebraic universes. As first proven in [18], we showed that the algebraic universes are sound with respect to the intuitionistic set theory **IZF**. Moreover, we saw that universes in Boolean categories of classes are sound with respect to the classical set theory **ZF**. In addition, we discussed the validity of the axiom of choice in the algebraic set theory setting. In particular, we saw that a well-known criterion from topos theory ensuring the validity of the axiom of choice translates into our framework after a slight modification. Thereafter, we investigated how statements in the internal logic of the outer category of classes relate to set theoretic statements about the algebraic universe that it contains. Roughly, we saw that constructions such as products, functions and powersets build in the outer category of classes can be squeezed into the algebraic universe via a coding procedure, delivering their set theoretic analogue in the universe. The squeezing results, as we like to call them, provided us with the major tool to transfer Tierney's forcing proof in the algebraic set theory setting. The possibility of such translations was made explicit in [6] and some examples were carried out in [30]. We further extended these results providing the necessary preparation for the next chapter.

As the key part of this thesis we showed how Cohen's model negating the Continuum Hypothesis fits into the algebraic set theory framework. The proof consisted of two steps. First, we observed that the sheaf theoretic proof given by Tierney in [29] can also be carried out in a category of classes, namely in the enlarged version of Tierney's sheaf topos. In this way we showed that the continuum hypothesis is falsified in the internal logic of the outer category of classes. In a second step, we used the squeezing results that we established in the previous chapter to prove that the statement in the internal logic of the category of classes can be translated to the algebraic universe that the category of classes contains. In this way, we obtained an algebraic model satisfying the usual **ZFC** axioms and the negation of the continuum hypothesis.

6.2 Discussion and Further Work

In this thesis we provided an algebraic model of set theory falsifying the continuum hypothesis, based on the topos theoretic result proved by Tierney [29]. Whereas the topos theoretic model only satisfies a bounded version of set theory, the setting of algebraic set theory allowed us to obtain a model satisfying the negation of the continuum hypotheses and all axioms of usual **ZFC** set theory.

Tierney's idea was to give an elementary proof using only constructions internal to a topos. In our presentation, however, we often made reference to some external universe of set theory. Especially in our considerations on large presheaves and large sheaves we used an external universe of sets that provided us a notion of smallness. So in particular, we assumed an external universe of set theory in our considerations in the last chapter.

But also in the context of algebraic set theory an elementary proof can be given. Van den Berg and Moerdijk [7] have shown that a category of classes is closed under the formation of sheaves. Using this result, one can show that Tierney's proof can be translated into the algebraic set theory framework without reference to an external universe of set theory.

Cohen's original method of forcing was soon recognized to be equivalent to an approach using Boolean valued models of set theory.[1] It would be interesting to see how the theory of Boolean valued models fits into the algebraic set theory framework. A natural starting point for such considerations could be a result by Higgs stating that the category of sheaves over a Boolean algebra and the category of Boolean valued models are equivalent categories.[2] The idea would be to translate Higgs result in the algebraic set theory setting, establishing the equivalence between the categories of large sheaves over Boolean algebras and

[1] See [4] for an account on Boolean valued models of set theory.

[2] A proof of Higgs' result can be found in [17] or [13].

the category of Boolean valued (class)-models. We expect that a Boolean valued universe of set theory then corresponds to the algebraic universe arising in the category of large sheaves over the Boolean algebra in question. It should be a small step to extend this result to the case of Heyting algebras.

Bibliography

[1] P. Aczel and M. Rathjen. "CST Book draft". 2010.

[2] S. Awodey and H. Forssell. "Algebraic models of intuitionistic theories of sets and classes". In: *Theory and Applications of Categories* 15 (2004), pp. 147–163.

[3] S. Awodey and M. A. Warren. "Predicative algebraic set theory". In: *Theory and Applications of Categories* 15 (2004), pp. 1–39.

[4] J.L. Bell. *Boolean-Valued Models and Independence Proofs*. Oxford University Press, 2011.

[5] B. van den Berg and I. Moerdijk. "Aspects of predicative Algebraic Set Theory I: Exact completion". In: *Annals of Pure and Applied Logic* 156 (2008), pp. 123–159.

[6] B. van den Berg and I. Moerdijk. "Aspects of predicative Algebraic Set Theory II: Realizability". In: *Theoretical Computer Science* 412 (2011), pp. 1916–1940.

[7] B. van den Berg and I. Moerdijk. "Aspects of predicative Algebraic Set Theory III: Sheaves". In: *Proceedings of the London Mathematical Society* (5) 105 (2012), pp. 1076–1122.

[8] B. van den Berg and I. Moerdijk. "Derived rules for predicative set theory: An application of sheaves". In: *Ann. Pure Appl. Logic* 163.10 (2012), pp. 1367–1383.

[9] F. Borceux. *Handbook of Categrical Algebra*. Cambridge University Press, 1994.

[10] M. C. Bunge. "Topos Theory and Souslin's Hypothesis". In: *Journal of Pure and Applied Algebra* 4 (1974), pp. 159–187.

[11] P. J. Cohen. "The Independence of the Continuum Hypothesis". In: *Proc Natl Acad Sci U.S. A.* 50(6) (1963), pp. 1143–1148.

[12] M. Fourman. "Sheaf models for set theory". In: *Pure Appl. Algebra* 19 (1980), pp. 91–101.

[13] M.P. Fourman and D.S. Scott. "Sheaves and logic". In: *Applications of Sheaves*. Ed. by M. Fourman, Ch. Mulvey, and D.S. Scott. Vol. 753. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 1979, pp. 302–401.

[14] P. Freyd. "The axiom of choice". In: *J. Pure Appl. Alg.* 19 (1980), pp. 103–125.

[15] K. Gödel. *The Consistency of the Continuum Hypothesis.* Princeton University Press, 1940.

[16] S. Hayashi. "On set theories in toposes". In: *Logic Symposia Hakone 1979, 1980.* Springer Berlin Heidelberg, 1981, pp. 23–29.

[17] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium, Vol. 1.* Clarendon Press, Oxford, 2002.

[18] A. Joyal and I. Moerdijk. *Algebraic Set Theory.* Cambridge University Press, Cambridge, 1995.

[19] K. Kunen. *Set theory–An Introduction to Independence Proofs.* Elsevier, 1980.

[20] S. Mac Lane and I. Moerdijk. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* Springer Verlag, 1992.

[21] M. Makkai and G. E. Reyes. *First Order Categorical Logic.* Ed. by A. Donald and B. Eckmann. Lecture Notes in Mathematics. Springer Verlag, 1977.

[22] J. Myhill. "Constructive set theory". In: *Journal of Symbolic Logic* 40 (1975), pp. 347–382.

[23] I. Moerdijk J. van Oosten. *Topos Theory, lecture notes.* 2007.

[24] J. van Oosten. *Basic Category Theory, lecture notes.* 2002.

[25] A. M. Pitts. "Handbook of Logic in Computer Science". In: ed. by D.M. Gabbay S. Abramsky and T.S.E. Maibaum. Vol. 5. Oxford University Press, 2000. Chap. Categorical Logic.

[26] A.K. Simpson T. Streicher S. Awodey C. Butz. *Relating topos theory and set theory via categories of classes.* Tech. rep. Carnegie Mellon University, 2003.

[27] D. S. Scott. "Category-theoretic models for intuitionistic set theory". manuscript slides of a talk given at Carnegie-Mellon University. 1995.

[28] A. K. Simpson. "Elementary axioms for the category of classes". In: *14th Annual IEEE Symposium in Logic and Computer Science.* 1999, pp. 77–85.

[29] M. Tierney. "Sheaf theory and the continuum hypothesis". In: *Toposes, Algebraic Geometry and Logic.* Springer Berlin Heidelberg, 1971.

[30] M. A. Warren. "Predicative Categories of Classes". MA thesis. Carnegie Mellon University, 2004.